



UNIVERSIDADE ESTADUAL DE CAMPINAS
INSTITUTO DE BIOLOGIA

Amanda Araujo Gomes Ferreira

“O PAPEL DE COUP-TFII NA AQUISIÇÃO DE DESTINOS
CELULARES DURANTE A DIFERENCIAÇÃO DE CÉLULAS-
TRONCO EMBRIONÁRIAS”

“THE ROLE OF COUP-TFII IN THE ACQUISITION OF
CELLULAR FATES DURING THE DIFFERENTIATION OF
EMBRYONIC STEM CELLS”

CAMPINAS

2017

AMANDA ARAUJO GOMES FERREIRA

**“O PAPEL DE COUP-TFII NA AQUISIÇÃO DE DESTINOS
CELULARES DURANTE A DIFERENCIAÇÃO DE CÉLULAS-TRONCO
EMBRIONÁRIAS”**

**“THE ROLE OF COUP-TFII IN THE ACQUISITION OF
CELLULAR FATES DURING THE DIFFERENTIATION OF
EMBRYONIC STEM CELLS”**

*Dissertação apresentada ao Instituto de
Biologia da Universidade Estadual de
Campinas como parte dos requisitos
exigidos para a obtenção do título de
Mestra em Genética e Biologia Molecular,
na área de Genética Animal e Evolução*

*Dissertation presented to the Institute of
Biology of the University of Campinas in
partial fulfillment of the requirements for the
degree of Master in Genetics and
Molecular Biology, in the area of Animal
Genetics and Evolution*

ESTE ARQUIVO DIGITAL CORRESPONDE À
VERSÃO FINAL DA DISSERTAÇÃO DEFENDIDA
PELA ALUNA AMANDA ARAUJO GOMES
FERREIRA ORIENTADA PELO DR. HENRIQUE
MARQUES BARBOSA DE SOUZA.

Orientador: DR. HENRIQUE MARQUES BARBOSA DE SOUZA

CAMPINAS

2017
FICHA CATALOGRÁFICA

Agência(s) de fomento e nº(s) de processo(s): FAPESP, 2015/06732-9

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Biologia
Mara Janaina de Oliveira - CRB 8/6972

Ar15p	<p>Araujo, Amanda Gomes Ferreira, 1992- O papel de Coup-TFII na aquisição de destinos celulares durante a diferenciação de células-tronco embrionárias / Amanda Araujo Gomes Ferreira. – Campinas, SP : [s.n.], 2017.</p> <p>Orientador: Henrique Marques Barbosa de Souza. Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Biologia.</p> <p>1. Fator II de transcrição COUP. 2. Células-tronco embrionárias. 3. Diferenciação celular. 4. Camadas germinativas. I. Marques-Souza, Henrique, 1977-. II. Universidade Estadual de Campinas. Instituto de Biologia. III. Título.</p>
-------	---

Informações para Biblioteca Digital

Título em outro idioma: The role of Coup-TFII in the acquisition of cellular fates during the differentiation of embryonic stem cells

Palavras-chave em inglês:

Coup transcription factor II

Embryonic stem cells

Cell differentiation

Germ layers

Área de concentração: Genética Animal e Evolução

Titulação: Mestra em Genética e Biologia Molecular

Banca examinadora:

Henrique Marques Barbosa de Souza [Orientador]

Cristina Pontes Vicente

Marilene Houmuth Lopes

Data de defesa: 05-05-2017

Programa de Pós-Graduação: Genética e Biologia Molecular

Campinas, 05 de Maio de 2017

COMISSÃO EXAMINADORA

Prof. Dr. Henrique Marques Barbosa de Souza

Profa. Dra. Cristina Pontes Vicente

Profa. Dra. Marilene Hohmuth Lopes

Os membros da Comissão Examinadora acima assinaram a Ata de Defesa, que se encontra no processo de vida acadêmica do aluno.

AGRADECIMENTOS

Agradeço primeiramente aos meus pais, Dilza e Helder, por terem sempre me apoiado e estado ao meu lado. Pelos momentos em que eu tive dúvidas ou me senti incapacitada e eles me ajudaram a ver que eu era capaz e me ajudaram a recuperar minhas forças. E por terem me ensinado a viver fazendo tudo com excelência mesmo que muitas vezes não reconhecido pelos demais. Por terem me ensinado valores, honestidade e a importância de ter pessoas amadas ao meu lado. E por terem me amado e me encorajado a fazer o que mesmo para mim parecia impossível. Aos meus pais eu devo tudo que conquistei e tudo que eu sou hoje, as pessoas que sempre estarão ao meu lado, e que eu sempre poderei contar.

Ao Henrique, meu orientador, eu gostaria de agradecer todos os ensinamentos. Foram 6,5 anos juntos, desde o início do laboratório até aqui. Onde nós precisamos enfrentar muitas dificuldades juntos. Ele sempre acreditou em mim, confiou e me mostrou que a minha idade ou o meu “status acadêmico” não faziam de mim uma pessoa menos capacitada do que as demais. Ele me mostrou o que é ser cientista, como fazer as perguntas científicas certas e como executar bem os experimentos para respondê-las. Ele me mostrou que não há limites, que por mais que os sonhos pareçam inalcançáveis nós podemos tentar alcançá-los. Ele também me ensinou como me relacionar com as pessoas e me mostrou diferentes maneiras de comunicar o meu trabalho. Além de todos os ensinamentos, ele foi um amigo que sempre se mostrou preocupado comigo e com o meu bem-estar a cima das atividades do laboratório.

À Viviane que trabalhou comigo no laboratório desde o início, se mostrando uma excelente colega de trabalho e amiga, muitas vezes me ajudando a erguer a cabeça e continuar. Ela também fez muitos experimentos essenciais que constam nessa dissertação, sempre com bom humor, coração aberto e sem cobrar nada em troca. Eu e a Viviane nos tornamos uma equipe, onde uma sempre ajudou a outra como se fossemos uma única pessoa. Agradeço por todos os anos juntas, por ela ser uma pessoa fácil de lidar e por conseguir manter o equilíbrio, a fluidez e a harmonia no nosso trabalho.

Aos demais colegas do laboratório, em especial: Mika, Diego, Thaísa e Lucimara. Por deixarem o ambiente mais agradável, por serem muito “team players” e sempre apoiarem uns aos outros e ajudar sempre da melhor maneira possível. Ao lado deles os dias no laboratório passam mais rápido e são dias mais leves. Agradeço por serem pessoas que me apoiaram muito e torceram muito pelo meu sucesso, assim como torço pelo sucesso deles. Também agradeço por terem me mostrado como é trabalhar em um ambiente onde não há competições internas ou pessoas querendo “puxar o tapete”, mas sim um ambiente onde se possível ajudarão a te “colocar de volta no tapete”. Tenho certeza que não importa para qual laboratório eu for, não encontrarei um time tão unido e agradável de trabalhar como esse.

À Ângela Saito, por ter me ensinado e ajudado muito durante esses anos, sempre com muita paciência e excelência, ela me ensinou como ser impecável na cultura de células.

À minha melhor amiga e companheira Nayara, por estar sempre ao meu lado tanto para discussões científicas quanto para assuntos não relacionados, me ajudando a me manter equilibrada. Por me contagiar com a sua alegria e fazer com o que os meus dias sejam mais positivos e esperançosos. Por caminhar comigo, comemorando as vitórias e me ajudando a reerguer nas derrotas. E por confiar e me fazer acreditar em um futuro melhor do que meus próprios olhos são capazes de ver.

Aos meus grandes amigos que moraram comigo (Rep Kaxota) e dividiram muito mais do que momentos mas a melhor parte da minha vida, em especial: Chip, Paulista, Bagaço, Buh, Chicken, Clits, Marina, Messi, Pegabem, Push, Barbe. Por me ajudarem e me apoiarem na minha carreira e também por me mostrarem um mundo novo onde eu pudesse me sentir amada e acolhida por ser quem eu sou. Por me mostrarem o verdadeiro significado da amizade e fazerem valer a pena cada momento juntos. Por me ensinarem que a diferença não nos faz melhor ou pior, mas nos faz únicos e dignos de admiração e respeito. E por inúmeras vezes terem despertado em mim o sentimento de que eu sou a pessoa mais sortuda do mundo por terem eles como amigos, não importa onde estivermos. Nossas histórias juntos são infinitas e fizeram com que eu pudesse crescer profissionalmente e pessoalmente, me transformando hoje em uma pessoa plena e feliz. Agradeço em especial meus eternos roomies Chip e Paulista, que a cima de tudo me mostraram que irmãos não são as pessoas que nascem dos mesmos pais, mas aquelas que sabemos que poderemos contar sempre em qualquer situação.

Junto as pessoas citadas a cima, agradeço também à aqueles que passaram pela minha vida e fizeram a diferença, me ensinaram e me ajudaram. A todas essas pessoas a minha gratidão por terem feito de mim quem hoje sou.

Por ultimo, agradeço à Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) pelo financiamento do projeto (processo 2015/06732-9).

RESUMO

Células-tronco embrionárias (CTE) têm ganhado grande destaque em pesquisas por serem o principal modelo *in vitro* para o estudo do desenvolvimento embrionário, da diferenciação de linhagens celulares específicas e da reprogramação de células somáticas em células pluripotentes induzidas iPS para o uso em terapia celular. Por serem capazes de auto-renovação e pluripotência, CTE têm sido amplamente utilizadas para o estudo do balanço entre o estado indiferenciado (pluripotente) destas células, e a determinação de destinos celulares específicos *in vitro*. O fator de transcrição Coup-TFII (Chicken Ovalbumin Upstream Promoter-transcription Factor II) possui um papel fundamental na regulação do desenvolvimento embrionário e na aquisição de destinos específicos. Em CTE humanas, Coup-TFII forma um circuito regulatório que controla a pluripotência, sendo importante para a saída destas células do estado de pluripotência direcionando-as para um destino neural. Visando caracterizar a função de Coup-TFII durante a saída do estado pluripotente e a aquisição de destinos celulares específicos em CTE murinas, a expressão de Coup-TFII foi modulada na linhagem E14-TG2A, a partir da superexpressão e da perda parcial, por silenciamento gênico, e total, por nocaute gênico pela técnica de CRISPR/Cas9. Estas células moduladas para Coup-TFII foram diferenciadas como corpos embrióides ou em camadas únicas de células, e a expressão e marcação de fatores de pluripotência e genes marcadores dos três folhetos germinativos foram avaliadas por PCR quantitativo e imunohistoquímica. Os resultados apontam para um atraso na saída da pluripotência nas células nocauteadas para Coup-TFII. Além disso, o estabelecimento da técnica de CRISPR/Cas9 por este projeto permitiu que a ferramenta de edição de genoma passasse a ser realizada rotineiramente no laboratório, gerando avanços importante nos projetos de pesquisa do grupo e permitindo o estabelecimento novas colaborações com outros laboratórios com interesse nesta tecnologia.

ABSTRACT

Embryonic Stem Cells (ESC) have been the an important focus of research mainly because of their potential as a model system to study the embryonic development, differentiation of specific cell lineages and somatic cell reprogramming into induced pluripotency cells (iPS) for cellular therapy. With the capacity of self-renewal and its pluripotency, ESC have been broadly used to study the balance between the undifferentiated state (pluripotent) and the determination of specific cell fates in vitro. The transcription factor Coup-TFII (Chicken Ovalbumin Upstream Promoter – Transcription Factor II) has a fundamental role in cell fate decision during embryonic development and has been recently shown to form a regulatory circuit that controls the pluripotency in human ESC, being important for the exit from pluripotency in these cells and directing them to a neural state. In order to characterize the function of Coup-TFII during the exit from the pluripotency state and acquisition of specific cell fates in mouse ESC, we attempted to modulate Coup-TFII expression in the E14-TG2A line through the over expression, gene silencing and gene knockout by CRISPR/Cas9 technique. These Coup-TFII modulated cells were differentiated in Embryoid Bodies or in monolayer of cells, and the expression of pluripotency and markers of the three germ layer were measured by qPCR and imunohistochemistry. The results indicate a delay in the exit from pluripotency in the cells that are Coup-TFII knockout. Moreover, the establishment of the CRISPR/Cas9 technique by this project allowed the routinely use of this genome editing tool in the laboratory, generating important advances in the group's projects and allowing the establishment of new collaborations with others laboratories with interest in this technology.

LISTA DE ABREVIATURAS

BSA: Albumina sérica bovina, do inglês bovine serum albumin

BRL: Células de fígado de rato bufalo, do inglês Buffalo rat liver cells

cDNA: molécula de DNA complementar, do inglês Complementary DNA

Cas9: Associado ao CRISPR 9, do inglês CRISPR associated 9

CE: Corpos Embrióides

ChIP-seq: Imunoprecipitação de cromatina seguida de sequenciamento massivo de DNA, do inglês Chromatin immunoprecipitation with massive parallel DNA sequencing

Coup-TFII: Chicken ovalbumin upstream promoter-transcription factor II

Coup-TFI: Chicken ovalbumin upstream promoter-transcription factor I

CTE: Células Tronco Embrionárias

CRISPR: Sequências curtas repetidas agrupadas e regularmente espaçadas, do inglês Clustered regularly interspaced short palindromic repeats

D: dia de diferenciação

DAPI: 4',6-diamidino-2-phenylindole

DBD: Domínio de ligação ao DNA, do inglês DNA Binding Domain

DEPC: Dietilpirocarbonato

DMEM: Meio modificado Dulbedo, do inglês Dulbecco's Modified Eagle Medium

E: Estádio embrionário

EDTA: Ácido etilenodiamino tetra-acético, do inglês Ethylenediamine tetraacetic acid

FACS: Seleccionador de células ativado por fluorescência, do inglês Fluorescent-Activated Cell Sorting

FBS: Soro Fetal Bovino, do inglês Fetal Bovine Serum

FoxA2: Forkhead Box A2

gDNA: DNA genômico

GFP: Proteína Fluorescente Verde, do inglês Green Fluorescent Protein

GMEM: Meio modificado Glasgow, do inglês Glasgow Modified Eagle's Medium

INFABIC: Instituto Nacional de Ciência e Tecnologia de Fotônica Aplicada à Biologia Celular

iPSC: Células-tronco pluripotentes induzidas, do inglês induced pluripotent stem cell

KLF4: Kruppel Like Factor 4

LBD: Domínio de ligação ao ligante, do inglês Ligant Binding Domain

LIF: fator inibidor de leucemia, do inglês Leukemia inhibitory fator

MEF: Fibroblastos extraídos de embriões de camundongos, do inglês Mouse Embryonic Fibroblast

miRNA/miR: microRNA

Oct4: Octamer-binding transcription factor 4

Pax6: Paired Box protein 6

pb: pares de base

PBS: tampão fosfato salino, do inglês phosphate buffered saline

PCR: Reação em cadeia da polimerase via transcriptase reversa, do inglês Reverse transcription polymerase chain reaction

PFA: paraformaldeído

pH: Potencial hidrogeniônico

Pop: População

qPCR: Reação em cadeia da polimerase via transcriptase reversa em tempo real, do inglês Quantitative Real-Time Reverse transcription polymerase chain reaction

RNA: Ácido ribonucleico, do inglês Ribonucleic acid

RPS29: Gene que codifica a proteína ribossomal 29

RRGP: Rede de Regulação Gênica da Pluripotência

sgRNA: Único guia de RNA, do ingles single guide RNA

shRNA: short hairpin RNA

Sox2: Sex determining region Y – Box 2

TV: Variante de transcrito, do inglês transcript variant

LISTA DE FIGURAS

Figura 1: Esquema mostrando as funções de COUP-TFII durante a organogênese de camundongos.....	20
Figura 2: Células passando pelo FACS, primeira separação, fluorescência verde.....	37
Figura 3: Células passando pelo FACS, segunda separação, viabilidade das células.....	38
Figura 4: Células passando pelo FACS, segunda separação, fluorescência verde.....	39
Figura 5: Comparação entre a porcentagem de fluorescência antes e depois da segunda separação e antes da terceira separação.....	40
Figura 6: Células passando pelo FACS, terceira separação, fluorescência verde.....	41
Figura 7: Microscopia de fluorescência evidenciando a diferença de expressão de GFP entre as populações 1 e 2.....	41
Figura 8: Imunohistoquímica para COUP-TFII em células transduzidas com vetores de shRNA e superexpressão.....	43
Figura 9: Citometria das diferentes condições testadas para lipofecção de células da linhagem E14TG2A.....	44
Figura 10: Gel de PCR para diagnóstico dos clones que sofreram microdeleção do exon 2 do COUP-TFII.....	46
Figura 11: Resultado do sequenciamento de um dos clones.....	46
Figura 12: Gel de PCR para diagnóstico dos clones que sofreram microdeleção do exon 1 e 2 do COUP-TFII.....	47
Figura 13: Resultado do sequenciamento dos clones para microdeleção dos exons 1 e 2.....	48

Figura 14: Gel de agarose evidenciando a presença de bandas extras após o tratamento com a enzima T7.....	49
Figura 15: Resultado do sequenciamento do clone 31 a partir do produto do PCR.....	50
Figura 16: Imunohistoquímica em CEs D6, utilizando anticorpo contra a proteína do Coup-TFII.....	51
Figura 17: Imunohistoquímica em CEs diferenciados por 4 dias.....	52
Figura 18: qPCR para o gene Oct4 em células selvagens e nocautes.....	52
Figura 19: Esquema representando a estratégia e os passos para o estabelecimento da linhagem repórter COUP-TFII-mCherry.....	53
Figura 20: Esquema mostrando os braços 5', 3' e mCherry clonados individualmente no vetor pBluescript.....	54
Figura 21: Gel de agarose mostrando os braços 3' e 5' após o PCR e purificação de cada um para a clonagem no vetor pBluescript.....	55
Figura 22: Gel de agarose mostrando o mCherry após o PCR e o vetor pBluescript linearizado.....	56
Figura 23: Gel de agarose para diagnosticar os clones que tiveram a inserção do 2a-mCherry no genoma.....	57
Figura 24: Resultado do sequenciamento de três clones mostrando que houve a inserção do p2a-mCherry.....	57
Figura 25: Gráficos gerados no citômetro para 8 filtros diferentes com o intuito de ver se há expressão de mCherry.....	59
Figura 26: Imunohistoquímica para COUP-TFII em CEs diferenciados por 10 dias a partir de células selvagens e células repórteres mCherry.....	60

LISTA DE TABELAS

Tabela 1: Lista com as sequências de todos os sgRNAs e Primers utilizados.....	30
---	----

SUMÁRIO

1. INTRODUÇÃO.....	14
2. OBJETIVOS.....	23
3. MATERIAIS E MÉTODOS.....	25
4. RESULTADOS E DISCUSSÃO.....	36
1) Modulação da expressão de Coup-TFII em linhagens de CTE por silenciamento e superexpressão a partir de vetores virais de expressão gênica.....	36
2) Ensaios de diferenciação por CE a partir de CTE moduladas para a expressão de COUP-TFII.....	42
3) Estabelecimento de CRISPR/Cas9 e geração de linhagens CTE nocaute para COUP-TFII.....	44
4) Ensaios de diferenciação celular com CTE nocaute para COUP-TFII.....	50
5) Estabelecimento de uma linhagem repórter COUP-TFII-mCherry por CRISPR/Cas9.....	53
6) Geração de linhagens de CTE nocaute para COUP-TFI por CRISPR/Cas9.....	61
5. CONCLUSÃO.....	62
6. REFERÊNCIAS.....	64
7. ANEXOS.....	69

1. INTRODUÇÃO

A vida animal se inicia com a fecundação de dois gametas haplóides para formar uma única célula diploide, o zigoto. Esta única célula passa por diversos processos precisos de clivagem e diferenciação que darão origem a um indivíduo composto por trilhões de células com mais de duzentos tipos celulares diferentes^{1, 2}. As primeiras células do embrião após as clivagens do zigoto, compondo a massa celular interna do blastocisto, são conhecidas por serem pluripotentes³, i.e., capazes de se diferenciarem em todas as células de um indivíduo adulto, inclusive nas células da linhagem germinativa. O conjunto de todas as transformações que ocorrem do zigoto até o nascimento e ainda no crescimento do indivíduo é chamado de desenvolvimento. Além dos fatores intrínsecos, estímulos externos⁴, epigenéticos⁵ e reorganizações morfogenéticas⁶ controlam a expressão diferencial de genes presentes no zigoto e em cada célula subsequente, de maneira com que se comportem diferentemente, mas reprodutivelmente, fazendo do desenvolvimento embrionário um processo preciso, dinâmico e eficaz⁴.

Assim como as células pluripotentes do embrião, existem células no organismo adulto conhecidas por serem multipotentes. Essas células são “indiferenciadas” e chamadas de células-tronco adultas^{7, 8}, por possuírem a habilidade de se auto-renovarem, isto é, de se proliferar em células-tronco semelhantes indefinidamente, e também a habilidade de se dividirem em células filhas específicas, ou diferenciadas, que possuem uma capacidade de proliferação limitada⁹. Essas células-tronco adultas são encontradas em tecidos que estão em constante renovação, como o sangue, epitélio do intestino e pele.

Células-tronco também são parte determinante na formação, manutenção e metástase de tumores. Assim como as demais células do tumor, as células-tronco tumorais acumulam mutações que as permitem proliferar e diferenciar-se em um nível anormal, mantendo também as características do tecido em que foram originadas¹⁰. Um tipo de câncer raro comumente encontrado nas gônadas é conhecido por teratoma, e a origem do nome “terato” do grego significa “monstro”. Este nome é dado por que esses tipos de tumores apresentam uma mistura “monstruosa” de tecidos adultos que pode conter dentes, pelos e até fragmentos ósseos¹¹. Por causa dessas características, esse tipo de tumor atraiu atenção e

curiosidade dos cientistas desde a antiguidade, sendo raramente encontrado e quase sempre nas gônadas. Por serem raros, os estudos de teratomas foram muito limitados até 1954, quando um grupo de pesquisadores relatou o desenvolvimento de teratomas nos testículos espontaneamente em 1% dos indivíduos da linhagem de camundongo 129¹². Mais tarde descobriram que a incidência destes tumores era muito maior em testículos de camundongos adultos enxertados com células de embrião¹³, o que possibilitou o estudo mais preciso desde o início da formação do teratoma.

Em 1964 um trabalho apresentou a ideia de pluripotência para os teratomas malignos, conhecidos por teratocarcinomas, quando uma única célula do tumor foi injetada no peritônio de um camundongo e deu origem a todos os diferentes tecidos encontrados em um teratocarcinoma¹⁴. Este fato demonstrou a existência de um tipo de células-tronco encontradas no teratocarcinoma, que são capazes de dividir indefinidamente, garantindo a malignidade do tumor e dando origem há vários tipos celulares do indivíduo adulto¹⁴. A constatação da pluripotência em células do teratocarcinoma e da promoção desses tumores através da enxertia de células retiradas de embriões gerou a hipótese de que teratocarcinomas seriam extremamente parecidos, e talvez idênticos, às células pluripotentes embrionárias¹¹. Teratocarcinomas passaram então a serem alvos de estudos que investigaram a possibilidade de se estabelecer linhagens de células pluripotentes para o estudo da diferenciação celular. Quando cultivadas isoladamente, estas células se mostraram muito heterogêneas^{15, 16}. No entanto, a partir da introdução de uma camada de células alimentadoras de fibroblastos de embrião na cultura, células do teratocarcinoma foram capazes de se manter indiferenciadas e proliferativas e, quando cultivadas em placas não aderentes, foram capazes de formar pequenos grumos que se diferenciavam em vários tipos celulares encontrados no embrião¹⁷. Estas esferas celulares se assemelhavam muito com embriões pós-implantação e por isso foram denominadas corpos embrióides, e passaram a ser considerados um excelente sistema modelo para o estudo da embriogênese e neoplasias¹⁸. Essas células extraídas dos teratocarcinomas de embriões que mantinham a pluripotência foram chamadas de células de carcinoma embrionário.

Uma vez relatada a semelhança entre os embriões de camundongos e as células de carcinoma embrionário, o próximo passo foi investigar se essas células eram capazes de contribuir para o desenvolvimento de embriões quando injetadas em blastocistos. Em 1974 foi reportado o primeiro sucesso na obtenção de animais quimeras saudáveis gerados de blastocistos injetados com células de carcinoma embrionário¹⁹. Porém estes estudos geraram muitos dados discrepantes entre os laboratórios, além de cariótipos anormais, o que dificultou a contribuição de células de carcinoma embrionário para a linhagem germinativa^{20, 21}. De qualquer forma, os estudos utilizando células de teratocarcinomas e células de carcinoma embrionário trouxeram muito conhecimento na forma e condições para o cultivo de células pluripotentes, além de terem ajudado a elucidar os antígenos de superfície presente nessas células e o estabelecimento de experimentos para testar a pluripotência das células^{11, 22}. Em consequência, esses estudos foram fundamentais para que cientistas conseguissem, em 1981, isolar e cultivar pela primeira vez células da massa celular interna de blastocistos maduros de camundongos, e posteriormente de humanos em 1998²³, trabalho que resultou em prêmio Nobel. Por serem isoladas do blastocisto e responsáveis pela formação do embrião completo *in vivo*, estas células foram denominadas Células-Tronco Embrionárias (CTE)^{24, 25}.

Com o isolamento das CTE, o estado de pluripotência celular, transiente durante o desenvolvimento embrionário, foi mantido por tempo indeterminado *in vitro*²⁴⁻²⁶. Para isso, estas células precisaram ser cultivadas sobre uma camada de fibroblastos extraídos de embriões de camundongo (MEF), mitoticamente inativos^{24, 25}. Além disso, o fator inibidor da leucemia (LIF) foi mostrado como essencial no meio de cultura para que essas células possam ser mantidas pluripotentes *in vitro*²⁷. Quando o LIF não é adicionado ao meio de cultura, CTE deixam o estado pluripotentes e iniciam suas diferenciações. Quando cultivadas em placas não aderentes na ausência de LIF, as CTE formam aglomerados celulares esféricos, denominados corpos embrióides, que, assim como as células de carcinoma embrionário, são um ótimo modelo para o estudo do desenvolvimento embrionário²⁸.

O estado de pluripotência é mantido na células através de uma rede de regulação gênica que garante que as células não interpretem sinais extrínsecos pró-diferenciação e mantenham o seu estado indiferenciado e proliferativo. Esta rede de

regulação gênica da pluripotência (RRGP)²² é principalmente coordenada por proteínas que ficaram conhecidas como fatores de pluripotência, como os genes OCT4 (Octamer-binding transcription factor 4)^{29, 30} e SOX2 (SRY-box 2) que, além de ser um fator de transcrição essencial para a pluripotência³¹, controla também a expressão de OCT4³². Quando SOX2 ou OCT4 são nocauteados, as células deixam o estado de pluripotência e ativam vias de diferenciação celular^{29, 31}. Contraintuitivamente, quando esses dois genes são superexpressos, ao invés de favorecerem o estado de pluripotência, esta expressão exacerbada destes fatores também promove a diferenciação celular³³. Isto demonstra que o balanço entre os fatores de transcrição da RRGp é extremamente preciso para que a pluripotência seja mantida. Apesar de dispensável para a manutenção de CTE³⁴, o gene NANOG também é considerado como um dos fatores de pluripotência essenciais para a RRGp, por causa de seu importante papel durante o estado de pluripotência das células da massa celular interna do blastocisto³⁵. Estes três genes controlando a RRGp (OCT4, SOX2 e NANOG) funcionam co-ocupando diversos promotores de outros genes, inclusive deles mesmos, em mecanismos de *feedback* e *feedforward* positivos para manter a pluripotência junto com outros fatores de transcrição, RNAs regulatórios, vias de sinalizações e mecanismos epigenéticos³⁶⁻³⁸. A pluripotência também é mantida pela ação dos micro-RNAs (miRNAs)³⁹, uma vez que o nocaute de genes essenciais para a biogênese dos miRNAs (Dicer e DGCR8) resultou na perda da capacidade das células de progredirem no ciclo-celular⁴⁰. Os grupos de miRNAs miR-290 e miR-302 promovem a auto-renovação e inibem a diferenciação celular⁴¹ através de mecanismos de silenciamento ou degradação de mRNAs de genes da diferenciação, mecanismos estes denominados de controles pós-transcricionais onde a tradução de tais genes é diminuída.

Um dos circuitos regulatórios que controlam a transição entre os estados de pluripotência e de diferenciação em CTE humanas envolve a regulação e ação repressora do fator de transcrição COUP-TFII (*Chicken Ovalbumin Upstream Promoter Transcription Factor II*, também conhecido como NR2F2 ou ARP-1). Durante o estado de pluripotência, OCT4 reprime diretamente a transcrição de Coup-TFII e ativa a expressão do miR-302, que por sua vez reprime COUP-TFII pós-transcricionalmente. No entanto, mediante indução da diferenciação, a expressão de COUP-TFII aumenta e reprime diretamente a expressão de OCT4, que deixa de

ativar a expressão do miR-302. Com a eliminação de seus dois repressores, Coup-TFII induz as células a se diferenciarem em precursores neurais^{42, 43}.

Juntamente com o seu parólogo COUP-TFI, COUP-TFII faz parte da subfamília COUP-TFs, e são membros da super família de receptores de hormônios esteroides^{44, 45}. Sem ligantes identificados para esses genes, eles são considerados receptores órfãos^{44, 46}. COUP-TFs se ligam aos mesmos sítios de ligação no DNA (DBD) que os receptores nucleares da Vitamina D, hormônios tireoides, ácido retinóico e retinóico X, competindo com a ação ativadora dessas moléculas, através de diferentes mecanismos de regulação⁴⁷. Esses genes são altamente conservados desde metazoários primitivo até humanos, sendo a subfamília de receptores nucleares mais bem conservada, com 96% de similaridade entre os DBD de camundongo e drosófila e 90% no domínio de ligação ao ligante (LBD), provavelmente sendo os membros primordiais da superfamília de receptores de hormônios esteroides^{48, 49}.

Entre eles, COUP-TFI e COUP-TFII compartilham 98% de homologia entre seus DBD⁴⁸. Mas apesar destes genes serem muito similares em sequência, eles estão em cromossomos diferentes e muitas das suas funções também são diferentes⁵⁰. Este trabalho focou nas funções de COUP-TFII, por que este gene já foi mostrado no circuito de regulação que controla a pluripotência com OCT4.

O COUP-TFII é um fator de transcrição que atua na ativação e repressão de diversos genes alvos^{45, 48, 50, 51}. COUP-TFII ainda promove a diferenciação celular, proliferação, migração, sobrevivência e comunicações intracelulares⁴⁵. Muitos trabalhos também já demonstraram funções regulatórias de COUP-TFII no desenvolvimento de órgão e tecidos como: cerebelo, cérebro, olho, coração, estômago, diafragma, braços e pernas, rim, adipose, testículos, vasos sanguíneos e linfáticos⁴⁵. Além de funções fisiológicas desempenhadas em indivíduos adultos⁴⁵, COUP-TFII também já foi mostrado facilitando a proliferação de metástases de tumores através do aumento da formação de vasos⁵².

Em camundongos, quando COUP-TFII é nocauteado em homozigose, os embriões são capazes de se desenvolverem até o estágio E9.5 apresentando anormalidades, e já no estágio E10.5 os homozigotos morrem com hemorragias no

cérebro e coração⁵³. Quando o nocaute ocorre em heterozigose, parte dos embriões consegue se desenvolver até o nascimento, também apresentando anormalidades, e dos que nasceram, 2/3 morre antes dos 21 dias de vida⁵³, evidenciando o papel essencial de COUP-TFII durante a embriogênese.

Outra evidência da importância de COUP-TFII na transição da pluripotência para a diferenciação é quando se faz o caminho oposto, a reprogramação de células somáticas para pluripotência. Células-tronco pluripotentes induzidas (iPSCs) foram descritas em 2006 por um grupo de pesquisadores que descobriu que forçar a expressão dos fatores de transcrição da RRPGE induz um estado de pluripotência em células somáticas⁵⁴. Os quatro fatores de transcrição que são suficientes para a reprogramação são: OCT4, SOX2, C-MYC e KLF4. A descoberta das iPSCs foi um grande avanço para o entendimento do desenvolvimento e medicina regenerativa, tornando possível a reprogramação de células de pacientes para a modelagem de suas doenças e para que estas sejam diferenciadas em outros tecidos para o tratamento de doenças e regeneração de tecidos paciente-específico⁵⁵⁻⁵⁷. O circuito entre OCT4, COUP-TFII e miR-302, que controla a pluripotência e diferenciação, atua também no retorno das células diferenciadas para o estado de pluripotência. Quando a reprogramação é feita com os quatro fatores de pluripotência conjuntamente com a expressão de miR-302 ou de um shRNA para silenciar o COUP-TFII, a eficiência da reprogramação celular aumenta, sendo reduzida quando a reprogramação é feita em condição de superexpressão de COUP-TFII⁴³.

Além do COUP-TFII ter este papel essencial na diferenciação, muitos trabalhos também já o descreveram como um gene determinante na ativação de redes de regulação gênica específicas para alguns tipos celulares, funcionando como um interruptor molecular que quando ativado muda totalmente o destino das células (Figura 1). Foi demonstrado que a expressão de COUP-TFII é essencial para que haja o desenvolvimento de veias durante a angiogênese. Quando COUP-TFII é nocauteado somente há a formação de vasos como artérias⁵⁸. Foi observado também que o papel de COUP-TFII como interruptor molecular, fez uma análise de imunoprecipitação de cromatina seguida por sequenciamento massivo (ChIP-seq) em corações de embriões de camundongos, demonstrando que COUP-TFII está amplamente ligado à genes que conferem a identidade de átrio às células. Quando

COUP-TFII é nocauteado em precursores cardíacos, o átrio perde sua identidade e se desenvolve similar a um ventrículo⁵⁹. Outro estudo muito importante no entendimento do COUP-TFII como um fator chave da diferenciação foi realizando com células mesenquimais. Células mesenquimais são células-tronco multipotentes que precedem quatro eventos: miogênese, condrogênese, osteogênese e adipogênese. Quando COUP-TFII é nocauteado, as células mesenquimais perdem a habilidade de gerar adipogênese e condrogênese e aumentam sua capacidade para osteogênese e miogênese⁶⁰.

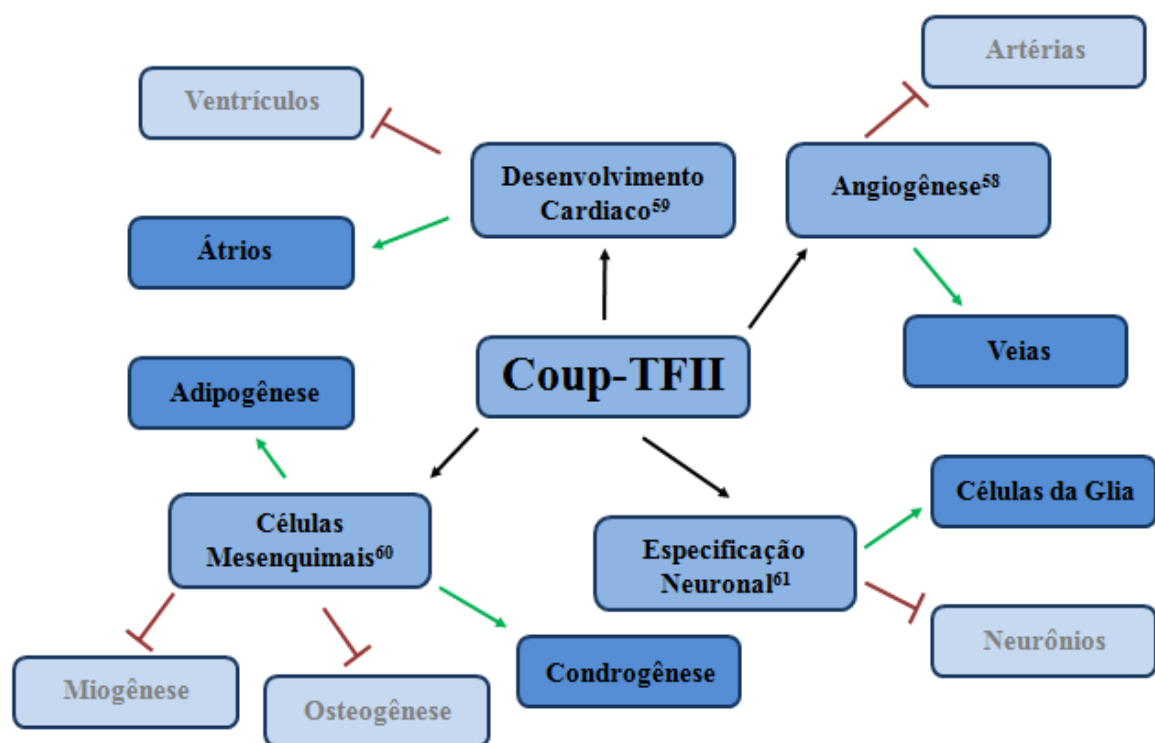


Figura 1: Esquema mostrando as funções de ativação e repressão de COUP-TFII durante a organogênese de camundongos.

Resultados preliminares do nosso laboratório com CTE mostraram, por citometria e imunohistoquímica, a co-localização de COUP-TFII com genes marcadores de ectoderme e mesoderme durante os estágios iniciais da diferenciação de CTE em Corpos Embrionários. O laboratório também demonstrou que a expressão de COUP-TFII é primeiramente detectada, por imunohistoquímica, após quatro dias de diferenciação, com um pico de expressão no 13º dia de diferenciação.

Por se tratar de um interruptor molecular expresso nos estágios iniciais da diferenciação de CTE e por estar co-localizado com marcadores de ectoderme e

mesoderme, a hipótese deste trabalho foi de que COUP-TFII estivesse funcionando como um interruptor molecular já nas primeiras tomadas de decisões de células pluripotentes. Este trabalho foi a primeira dissertação do laboratório focada nos aspectos moleculares da ação de Coup-TFII em CTE, e representa a base de praticamente todas as técnicas e ferramentas sendo utilizadas no laboratório até o momento dessa dissertação para o estudo da função gênica em Células-Tronco Embrionárias.

O objetivo deste trabalho foi estudar as funções de Coup-TFII durante o início do desenvolvimento através da modulação da sua expressão. Primeiramente a estratégia utilizada foi a de modular a expressão de Coup-TFII através de vetores de silenciamento (shRNA) e de superexpressão. Apesar de estabelecermos linhagens estáveis com os vetores de expressão para a modulação da expressão de Coup-TFII⁶¹, concluímos que tais vetores foram incapazes de silenciar ou superexpressar Coup-TFII. Portanto, foi estabelecida e utilizada a técnica de CRISPR/Cas9. A técnica de edição genômica CRISPR/Cas9 foi amplamente aplicada para o estudo das funções gênicas⁶². Esta técnica permite nocautear genes ou inserir novos genes ao genoma com a introdução de vetores expressando transientemente a enzima Cas9 e guias de RNA que a direcionam para fragmentar o DNA genômico em regiões específicas⁶². Portanto, a segunda etapa do trabalho sendo descrito nesta dissertação focou em estabelecer CRISPR/Cas9 no laboratório visando gerar linhagens de CTE nocauteadas para a expressão de Coup-TFII e linhagens celulares que contenham um gene repórter no mesmo cassete de expressão do Coup-TFII.

Um dos desafios da geração de nocautes gênicos por CRISPR/Cas9 foi a existência de variantes de transcritos e alternativos de splicing. No caso de COUP-TFII, o gene pode ser transcrito a partir de dois promotores diferentes, gerando a Variante de Transcrito 1 (TV1) e a Variante de Transcrito 2 (TV2). As duas variantes de transcrito compartilham o segundo e o terceiro exon do gene, sendo que o primeiro exon da TV1 contém o DBD da proteína e o primeiro exon da TV2 não. Sendo assim, apenas a TV1 é capaz de se ligar em promotores de outros genes para os regular e a TV2 foi mostrada funcionando como inibidor da TV1 quando em altas concentrações, por se ligar com a TV1 e impedir que esta se ligue aos

promotores dos genes que regula⁶³. Ademais, a partir de trabalhos já existentes é possível concluir que a deleção de qualquer um dos exons da TV1 altera a capacidade do COUP-TFII de regular outros genes^{51, 53}. Sendo assim, para o nocaute de COUP-TFII utilizando CRISPR/Cas9, este trabalho focou em três estratégias diferentes: a deleção apenas o exon 2 (que é comum para TV1 e TV2 e compõe parte do LBD); a deleção dos exons 1 da TV1 (que contém o DBD) e do exon 2; e o frameshift da proteína COUP-TFII-TV1 inteira a partir do exon 1 da TV1.

2. OBJETIVOS

Estudar o papel de COUP-TFII durante a transição entre a pluripotência e a diferenciação celular de CTE murinas.

Objetivos específicos:

- 1) Modulação da expressão de Coup-TFII em linhagens de CTE por silenciamento e superexpressão a partir de vetores virais de expressão gênica.

Transfectar CTE selvagens da linhagem E14-TG2A com vetores de silenciamento (shRNA) e de superexpressão da região codante de COUP-TFII e selecionar estas células por Citometria de fluxo ou seleção por resistência à antibiótico.

- 2) Ensaio de diferenciação a partir de CTE moduladas para a expressão de COUP-TFII.

Diferenciar CTE moduladas para Coup-TFII por CE e verificar a expressão por qPCR e co-localização por imunohistoquímica a fim de analisar se a modulação de COUP-TFII altera o padrão de expressão de genes marcadores dos três folhetos germinativos: Pax6/Nestin (ectoderme), Brachyury T (mesoderme) e Gata4/FoxA2 (endoderme).

- 3) Estabelecimento de CRISPR/Cas9 e geração de linhagens CTE nocaute para COUP-TFII.

Estabelecer a técnica de CRISPR/Cas9 no laboratório e gerar diferentes linhagens nocautes para COUP-TFII.

- 4) Ensaio de diferenciação celular com CTE nocaute para COUP-TFII

Diferenciar CTE e analisar se o nocaute de COUP-TFII altera a expressão de genes marcadores dos três folhetos germinativos.

- 5) Estabelecimento de uma linhagem repórter COUP-TFII-mCherry por CRISPR/Cas9

Estabelecer linhagens que expressem o gene da proteína fluorescente mCherry sobre o controle da região genômica de COUP-TFII for expresso.

6) Geração de linhagens de CTE nocaute para COUP-TFI por CRISPR/Cas9

Estabelecer linhagem de CTE nocautes para COUP-TFI utilizando a técnica de CRISPR/Cas9 para tirar o gene do frame de leitura (frameshift).

3. MATERIAIS E MÉTODOS

Cultivo e manutenção das células

CTE murinas da linhagem E14-TG2A cedidas por Dr. Joshua Brickman (Centre Development in Stem Cell Biology, Institute for Stem Cell Research, School of Biological Sciences, University of Edinburgh) foram cultivadas indiferenciadas em meio GMEM (Sigma, USA) contendo 10^3 U/mL do fator inibidor de leucemia (LIF) (ESGRO-LIF; Gibco), 0,1 mM B-mercaptoetanol, 1% v/v de aminoácidos não essenciais, 0,5% de Penicilina/Estreptomicina, 2 nM L-Glutamina, 1% de piruvato de sódio e 15% soro fetal bovino (FBS - Gibco), na ausência de células alimentadoras (*feeder cells*). Esta linhagem não precisaram das células alimentadoras por que foram isoladas em um meio que contém células de fígado de rato búfalo (Buffalo rat liver cell – BRL)⁶⁴ e depois cultivadas em placas gelatinizadas.

A cada dois dias, as células foram transferidas para outra placa, para que a confluência atingida não seja tão alta, pois a confluência acima de 80% leva as células a se diferenciarem ou morrerem. Então, o meio das células era retirado, as células eram lavadas com 4mL de PBS 1X e tripsinizadas com Tryple Express (Gibco) 1: 1 PBS 1X. Em seguida, em uma garrafa T25 pré-gelatinizada com gelatina 0.1%, foram colocadas aproximadamente 1 milhão de células, com 5mL de meio.

Transdução viral das células

Para a transdução das células com vírus, os vetores doados pelo Dr. Hideyuki Okano foram usados para a montagem de lentivirus pelo Laboratório de Vetores Virais do Dr. Marcio Chaim Bajgelman, do LNBio.

Os vetores utilizados foram:

1. Um vetor que expressava uma molécula de RNA complementar ao mRNA do Coup-TFII com uma sequência invertida e complementar a ela, separadas por uma sequência inerte, que resultou na transcrição de uma molécula de RNA que se auto-complementou e formou uma estrutura conhecida como short hairpin (shCoup-TFII), promoveu o

silenciamento de Coup-TFII. Este vetor expressava concomitantemente o gene que codifica a proteína fluorescente verde (GFP).

2. Um vetor que expressava uma molécula RNA que também resultou em um RNA short hairpin, mas sem complementariedade a algum mRNA da célula (shControle), servindo como controle à introdução de moléculas hairpin na célula. Este vetor também promoveu a expressão do gene da proteína fluorescente verde (GFP).
3. Um vetor que expressava a região codante do gene Coup-TFII, resultando em uma superexpressão do Coup-TFII exógena. Junto com a expressão do Coup-TFII exógeno, também foi expresso um gene de resistência a higromicina.

No dia anterior à transdução viral, 80 mil células foram plaqueadas em poços de placa de 24 wells pré-gelatinizados. No dia da transdução, os vírus foram colocados nas células em uma proporção de 1:1, um vírus por célula. No dia seguinte, o meio foi trocado e as células foram lavadas com 500 uL de PBS 1X. Após 48 horas da transdução, já foi possível ver as células verdes, expressando o gene da proteína fluorescente GFP, que foi expresso no mesmo vetor.

Seleção - FACS

Visando a utilização apenas de células que foram transduzidas e expressam os vetores, foi realizada uma seleção baseada no gene repórter GFP (vetores shCoup-TFII e shControle), que era expresso constitutivamente por esses vetores. Essa seleção das células foi realizada via FACS (Fluorescence-Activated Cell Sorting), que nos possibilitou separar as células também por sua intensidade de fluorescência verde. As células tripsinizadas, foram lavadas com PBS 1X e centrifugadas duas vezes a 1300 rpm por 5 minutos, para retirar o excesso de meio, que pode influenciar na captação da fluorescência. As células então foram separadas pelo BD FACS ARIA™ Ilu do Hemocentro da UNICAMP, e foram readeridas em uma placa gelatinizada.

O processo de separação não foi 100% preciso, então foi necessário repetir o procedimento para aumentar a porcentagem de células expressando GFP, i. e., contendo os vetores de shCoup-TFII ou shControle.

Seleção – Higromicina

As células transduzidas com o vetor de superexpressão de Coup-TFII foram tratadas com 200 ug/mL de higromicina B (Invitrogen) adicionada ao meio e trocada todos os dias por aproximadamente 5 dias, quando todas as células do controle morreram.

Ensaio de diferenciação celular

Culturas de CTE foram induzidas à diferenciação a partir da formação de corpos embrióides (CE), pela técnica de gotas penduradas (do inglês hanging drop), utilizando o meio de cultura DMEM (Gibco), 1% v/v de aminoácidos não essenciais, 2 nM L-Glutamina, 1% de Penicilina/Estreptomicina, 0,1 mM B-mercaptoetanol e 20% soro fetal bovino (FBS) sem a presença de LIF, como descrito em Siegle et al, 1997⁶⁵. Gotas de 20uL foram colocadas na tampa de uma placa de petri contendo 10mL de água. Após três dias nas gotas, os CEs foram coletados em placas de petri com fundo de Agar para que não haja adesão ao fundo da placa, a cada dois dias na placa, metade do meio foi trocado.

As células também foram diferenciadas em camadas únicas (*monolayer*), onde foram plaqueadas em placa de 24 poços utilizando mil células por poço e o mesmo meio citado à cima para os Corpos Embrióides.

Extração de RNA

O RNA total dos CE ou células em camadas únicas em diferentes dias de diferenciação foi extraído utilizando-se o reagente Trizol (Invitrogen) seguindo o protocolo do fabricante. As amostras foram maceradas em eppendorfs contendo 1mL de Trizol e incubados por 5 min à temperatura ambiente. Em seguida, adicionou-se 200 µL de clorofórmio a cada amostra, agitando por 15 s e incubando novamente por 3 min à temperatura ambiente. Na sequência, o material foi

centrifugado a 14000 rpm por 15 min a 4°C. Após a centrifugação, a fase aquosa foi removida, então foi adicionado 500 µL de isopropanol e incubado por 10 min à temperatura ambiente. O material foi centrifugado a 14000 rpm por 10 min a 4°C para formação do pellet na parede do tubo. Após essa etapa, o sobrenadante foi removido, e adicionado álcool etílico 70% a 4°C e o material foi então novamente centrifugado a 7000 rpm por 5 min a 4°C. Finalmente, o sobrenadante foi retirado e o pellet permaneceu durante 15 min exposto para secar. Em seguida, o material foi ressuspensionado em 50 µL de água DEPC (*Diethyl pyrocarbonate*).

A avaliação da integridade das amostras de RNA total foi feita por meio da eletroforese em gel de agarose a 1% na cuba horizontal, com corrente contínua a 100V durante 40 min e a concentração e pureza do RNA foi analisada utilizando o espectrofotômetro Nanodrop ND-1000 UV-Vis (Nanodrop Technologies), a partir da absorbância a 260 nm, e razões 280/260 nm e 260/230 nm, respectivamente.

Para síntese de cDNA foi utilizado o kit RevertAid H Minus First Strand cDNA Synthesis Kit (Fermentas), a partir de amostras de RNA total. A síntese de cDNA foi realizada a partir de 1 µg de RNA total, acrescido de 1 µL de Primer Oligo-(dT)18 (100uM) e quantidade suficiente de água DEPC para um volume final de 12 µL. Essa mistura foi incubada à 5 min por 65°C para desnaturação. Após este período de incubação, foram adicionados 4 µL de Tampão de Reação (5X), 1 µL de dNTP (10mM), 1 µL de Inibidor de RNase Ribolok (20 U/µL) e 1 µL de RivertAid H Minus Transcriptase Reversa (200 U/µL). As amostras foram incubadas novamente por 1 hora a 42°C e 5 min a 70°C e o cDNA originado foi quantificado por espectrofotometria no Nanodrop, como descrito anteriormente.

qPCR

A técnica de PCR quantitativo foi empregada para validar os dados obtidos através do transcriptoma diferencial entre as células expressando e não expressando Coup-TFII. A eficiência de amplificação (E) de cada um dos pares de primers foi determinada por meio da confecção de uma curva padrão de cinco pontos, gerada a partir de reações em duplicatas de qPCR nas quais foram utilizados diluições seriadas (1, 1:10, 1:100, 1:1000, 1:10000) de cDNA de grupos onde cada gene apresenta alto nível de expressão de acordo com dados da literatura.

Os genes de referência utilizados como controle endógeno foram Rps29 e Beta-Actina (Tabela 1). Os resultados foram normalizados usando os valores de Ct (cycle threshold) desses genes de referência. Os genes analisados por essa técnica foram: Coup-TFII, Brachyury, Coup-TFI, Gata4, Pax6, Oct4, Sox2, FoxA2 e Nestin (Tabela 1).

As reações de qPCR foram realizadas com 200 ng de cDNA de CTE indiferenciadas e CEs em diferentes estágios de diferenciação. As reações de qPCR foram feitas em triplicatas biológicas e triplicatas técnicas no equipamento 7500 Fast Real-Time PCR System (Applied Biosystems) com o kit SYBR Green JumpStart Taq ReadyMix (Sigma-Aldrich). Para um volume final de 10 μ L de reação foram adicionados 5 μ L de SYBR® Green Master Mix com ROX, 0,25 μ L de primer forward (10 μ M), 0,25 μ L de primer reverse (10 μ M) e H₂O tratada com DEPC para 10 μ L. As reações foram então adicionadas em placas ópticas MicroAmp Fast Optical 96wells Reaction Plate with Barcode (0.1 mL) (Applied Biosystems®). Uma vez preparadas as amostras, as reações foram realizadas com as seguintes condições de ciclagem: uma etapa inicial de desnaturação a 95°C por 10min, 40 ciclos de 95°C por 1min, 60°C por 15s e 72°C por 20s, seguindo de uma etapa final para a construção da curva de melting de 72°C por 20s, 60°C por 1min e 95°C por 15s.

Os resultados gerados foram analisados por delta-delta ct. Onde primeiramente é calculado uma diferença (delta1) entre a expressão do gene analisado e a expressão do gene endógeno da mesma amostra para a amostra controle e para as outras. Depois é calculado a diferença (delta2) entre o delta1 das amostras analisadas e o delta1 da amostra controle. Esse delta2 tem seu valor negativo elevado ao quadrado para o cálculo de quantas vezes (fold) a expressão da amostra analisada é maior ou menor do que a expressão da amostra controle. Depois um teste estatístico ANOVA seguido de Tukey é aplicado através do programa GraphPad Prism para analisar se há diferenças significativas de expressão gênica entre os grupos analisados.

sgRNAs	Sequência 5' - 3'	Primers PCR/Seq	Sequência 5' - 3'	Primers qPCR	Sequência 5' - 3'
sgRNA 1 F	CACCGCCCGGATGAGGGTTTCGATG	21	CTAGTCTAGATTCTCTAAAGATAACAGGGAGGG	Coup-TFII F	GTGTGCTTTGGAAGAGTACGTTAG
sgRNA 1 R	AAACCATCGAAACCCCTCATCCGGGC	23	GATCGGATCCTTGAATTGCCATATATGCCAG	Coup-TFII R	CAATTGCTCTATGACTGAGGAGGA
sgRNA 3 F	CACCGCAGCAGCTGGCGGACCCCC	24	TCGAATCGATAGGGGAGTGAAACAGAGAAAG	Brachyury F	ACCCAGACTCGCCCAATTTT
sgRNA 3 R	AAACGGGGGTGCGGCCACGTGCTGC	25	GTACGGTACCAGTTTACCTATAAGTGCCAAACAGT	Brachyury R	ACATGATCTGTCCCTCCCA
sgRNA 4 F	CACCGCGTCCGGACTTCTGCTCCCC	34	GATCGGATCCGAGCCACGAACCT	Coup-TFI F	ATTCTTCTCGCTGAACCGC
sgRNA 4 R	AAACGGGGAGCAGAAGTCCGGACGC	35	TCGAATCGATCTTGTACAGCTCGTCCATGC	Coup-TFI R	GCCAAAGGAAGTGTCCAT
sgRNA 11 F	CACCGTGTCCGCGCAGCAGTTTAAAC	92	AAACCAACAAAGCTGAGCCG	Gata4 F	TGTGCCAACTGCCAGACTA
sgRNA 11 R	AAACGTTAAACGTGCTGCCGACAC	93	AAATCCCCTCCGACTGCAAC	Gata4 R	TTTGAATCCCTCCTTCCGC
sgRNA 15 F	CACCGCCACAGTTAATGATCAGCCT	98	CCAACCTCAAATCAACTAGCCC	Pax6 F	GCACATGCAAAACACATGA
sgRNA 15 R	AAACAGGCTGATCATTAACTGTGGC	102	GGCCATCTGGGTAAACCTTAAT	Pax6 R	ACTTGAGCGGAAGTGCAC
sgRNA 16 F	CACCGCCTCCTGCAGCTGTACAGAG	103	ATGGTAGTCTTTAGCCTCTCCT	Oct4 F	AGAGGATCACCTTGGGGTACA
sgRNA 16 R	AAACCTCTGTACAGCTGCAGGAGGC	107	TGCTTCAAAATTGCCTGAAGTGT	Oct4 R	CGAAGCGACAGATGGTGGTC
sgRNA 17 F	CACCGTCAAAGTGCACTCCATCTT	108	GTTTCGGAAGCTCTTGCTTCG	Sox2 F	AGCAATGGTCTTATGTTGGAGC
sgRNA 17 R	AAACCAAGTATGGAGTGCACTTTGAC	111	TTACAGTCCATTAAATTTCTCCCC	Sox2 R	TGCTTTGTCCGTATCCAGTGC
sgRNA 18 F	CACCGTGGGAACATGATATATCTGC	B-Actina F	CAGGGTGTGATGGTGGGAAT	FoxA2 F	GGACCCCAAGACATACCGAC
sgRNA 18 R	AAACGCAGATATATCATGTTCCAC	B-Actina R	GCTCAGTAACAGTCCGCCT	FoxA2 R	ATCTTGTGGGCTCTGCTG
sgRNA 24 F	CACCGAGAAAGAAAAGGCAAAAGAC	U6 Forward	GAGGGCCTATTTCCTCATGATTCC	Nestin F	GTCCCTTAGTCTGGAAGTGGC
sgRNA 24 R	AAACGTCTTTTGCCTTTTCTTTCTc			Nestin R	TTGGGGTCAGGAAAGCCAAG
sgRNA 26 F	CACCGTAAAAAGAAAGAGAGAAAAG			B-Actina F	GTGGATCAGCAAGCAGGAGT
sgRNA 26 R	AAACCTTTTCTCTCTTTTCTTTTAc			B-Actina R	AGGGGTGTAACCGCAGCTCA
sgRNA 27 F	CACCGCTCTCACAACCCAGCCCAATCC			RPS29 F	GGGCGCTGAAGGCAAGATGGG
sgRNA 27 R	AAACGGATTGGGCTGGGTTGGAGGc			RPS29 R	TTGGAGCAGACGCGGCAAGAG

Tabela 1: sequência dos sgRNAs, primers para RT-PCR, sequenciamento e qPCR.

Imunohistoquímica

As reações de imunohistoquímica foram realizadas com CE em suspensão ou em células diferenciadas em camadas únicas sobre lamínulas, fixados com PFA 4% por 15 minutos e permeabilizados com Triton 0,5% diluído em PBS 0,1 M, pH 6,8, por 30 minutos. Os CE ou células em camadas únicas foram incubados em solução tampão de bloqueio contendo Triton 0,2%, glicina 0,3% e BSA 3%, diluídos em PBS 0,1M, pH 6,8, para bloqueio das reações inespecíficas, por 30 minutos à temperatura ambiente. Para as reações de imunohistoquímica, as amostras foram incubadas por 16 horas à 4°C com os anticorpos primários específicos para detecção de Coup-TFII (ab41859), Sox2 (sc17319) e Oct4 (sc9081). Após lavagem com PBS 0,1 M pH 6,8, os CE foram incubados durante 3 horas na temperatura ambiente com o anticorpo secundário específico e então lavados e incubados com DAPI, para evidenciar núcleos, na concentração de 1:1000, ambas incubações de 1 hora.

As imagens de CE em 3D ou de células em camadas únicas foram obtidas com o microscópio Confocal LSM780 – NLO INFABiC-UNICAMP. As imagens relatadas representaram uma projeção do eixo Z dos corpos embrióides íntegros (secção ótica).

CRISPR/Cas9

A edição do genoma por CRISPR/Cas9 foi realizada utilizando-se o vetor px462 da Addgene. Esse vetor contém a endonuclease Cas9 mutada, conhecida por Cas9nickase (CRISPR/Cas9n) que ao invés de cortar as duas fitas do DNA (como a

Cas9) tem a capacidade de cortar apenas uma, e então precisa ser utilizado um par de guias para que o efeito da técnica seja atingido. A utilização dessa endonuclease mutada aumentou a especificidade da edição, uma vez que esta mutação diminuiu muito as chances de que outra região do genoma sofresse o corte que não a região que se deseja cortar. O vetor px462 ainda possuía um promotor U6 para a transcrição dos RNAs guias que foram inseridos no vetor utilizando um sítio para a enzima de restrição BbsI. Além da Cas9n e do guia, o vetor também possuía um gene de resistência a puromicina.

- **Estratégias escolhidas para as edições**

Para a edição do COUP-TFII foram desenhadas as seguintes estratégias:

- i. Microdeleção do exon 2 do COUP-TFII.

Para isso foram desenhados um par de guias antes do exon 2 (sgRNA 15 e 16 – Tabela 1) e um par de guias depois do exon 2 (sgRNA 17 e 18 – Tabela 1).

- ii. Microdeleção dos exons 1 e 2 do TV1 do COUP-TFII.

Para isso foram desenhados um par de guias antes do exon 1 (sgRNA 3 e 4 – Tabela 1) e utilizado o mesmo par de guias depois do exon 2 da estratégia anterior.

- iii. Frameshift a partir do exon 1 do COUP-TFII TV1.

Para isso foi desenhado um par de guias no início do exon 1 (sgRNA 3 e 4 – Tabela 1).

- iv. Inserção de gene repórter mCherry no final do COUP-TFII.

Para isso foi desenhado um par de guias anelando no final do exon 3 (sgRNA 1 e 11 – Tabela 1) do COUP-TFII e foi também desenhado um vetor doador de homologia que continha a sequência codante do aminoácido auto-clivante 2A, seguido pela sequência codante do gene repórter mCherry, flanqueado por duas regiões homologas ao COUP-TFII, inserindo o gene mCherry no lugar do sítio de terminação de Coup-TFII.

Nocaute do parálogo COUP-TFI

- v. Frameshift do exon 2 do COUP-TFI.

Para isso foram desenhados um par de guias anelando no início do exon2 (sgRNA 26 e 27 – Tabela 1) do COUP-TFI.

- **Construção dos sgRNAs**

Para construir os sgRNAs foram desenhados oligos que contivessem 20 nucleotídeos complementares à região de interesse de edição, que fosse adjacente à uma sequência específica de nucleotídeos: NGG. Essa pequena sequência é essencial para que a Cas9n faça o corte no local certo, diminuindo a chance de cortes em outros locais do genoma caso o guia pareie de forma não específica. Além disso, os oligos foram desenhados contendo um sítio de clivagem para a enzima de restrição BbsI, permitindo a inserção dos guias no vetor px462.

Para cada sgRNA foram desenhados e sintetizados dois oligos de fita simples complementares. Os dois oligos foram ressuspensos a uma concentração final de 100 mM e então 1 uL de cada oligo foi misturado com 1 uL de T4 PNK Buffer 10x (New England Biolabs), 1 uL da enzima T4 PNK 10000U/mL (New England Biolabs), 0.5 uL de ATP 25mM (Thermo) e 5.5 uL de água depc. Essa mistura foi incubada a 37°C por 30 minutos e depois a 95°C por 5 minutos e chegou a 25°C descendo 5°/minuto. Após esses ciclos, os dois oligos encontravam-se anelados em dupla fita.

O oligo dupla fita foi então inserido no vetor px462, que possui um sítio BbsI para inserção dos guias. Utilizando-se 10 ciclos que intercalam digestão e ligação. Primeiramente os oligos foram diluídos 1:200 e então 2 uL foram adicionados a uma mistura que continha 1 uL do vetor 100 ng/uL, 4 uL do tampão Tango 10X (Thermo), 1 uL de DTT 10mM (Thermo), 0.5 uL de ATP 25mM (Thermo), 3 uL de BbsI 5000U/mL (New England Biolabs), 0.5 uL de enzima T4 ligase (Thermo) e 28 uL de água depc. Os ciclos de digestão foram de 37°C por 5 minutos e os ciclos de ligação foram de 22°C por 5 minutos. Esses sgRNAs foram então transformados em bactérias e sequenciados utilizando o primer U6 Forward (Tabela 1). Daqui por diante o conjunto de um guia clonado em um vetor px462 foi denominado sgRNA, seguido pelo número que representa o referido guia.

- **Lipofecção dos sgRNAs**

No dia anterior à lipofecção, 1 milhão de células foram plaqueadas em poços da placa de 6 poços gelatinizados com gelatina 0.1% em meio de cultura padrão sem penicilina e streptomicina (P/S). No dia seguinte, faltando uma hora para adicionar a lipofectamina às células, o meio das células foi retirado e foi adicionado à cada poço 500 uL de Opti-MEM (Gibco) sem suplementação. Em 150 uL de Opti-MEM foi diluído 7.5 uL de Lipofectamina 2000 (Invitrogen) e em outros 150 uL de Opti-MEM foram misturados os vetores em concentrações iguais para uma concentração final total de 3 ug de vetor. Após 5 minutos de repouso, os meios contendo lipofectamina e vetor foram misturados e incubados à temperatura ambiente por 30 minutos. Após a incubação, os 300 uL de Opti-MEM contendo lipofectamina e vetores foram adicionado às células totalizando 800 uL de Opti-MEM por poço. Um poço recebeu apenas vetor GFP e lipofectamina para controle. Após 4h da adição de lipofectamina às células, o meio foi totalmente retirado e substituído por meio padrão de cultura sem a presença dos antibióticos P/S.

Após 24 horas da retirada da lipofectamina das células, foi retirado o meio e adicionado novo meio sem P/S contendo 1 ug/mL de Puromicina, já que os vetores contendo os sgRNAs expressam o gene de resistência à puromicina. Após 48 horas da retirada da lipofectamina foi retirado o meio e adicionado novo meio sem P/S contendo, 1ug/mL de Puromicina. Após 72 horas da retirada da lipofectamina as células foram lavadas duas vezes com PBS 1x para retirar todo o resto de puromicina e foi adicionado meio padrão de cultura de células. Após 5 dias da retirada da lipofectamina os clones foram separados.

- **Separação dos clones**

A edição do genoma celular é um evento que ocorre independentemente em cada célula em cultura. Após a seleção com puromicina, somente as células que receberam pelo menos um dos sgRNAs sobrevivem. Entretanto, faz-se necessário a separação de clones para gerar populações celulares que possuem a mesma edição. Alguns eventos possíveis de acontecer nas células resistentes à puromicina são: (1) células em que apenas um dos guias entrou e a edição não resultaria na deleção desejada, (2) células em que os dois guias entraram mas não houve corte e estas continuam sendo selvagens, (3) células onde houveram os cortes mas o resultado não é o esperado e (4) células que receberam os guias e as deleções

desejadas aconteceram. Destas, cada evento bem sucedido pode ocorrer de forma distinta em cada célula. Por estes motivos, as células precisam ser separadas uma a uma para que possam crescer novamente e todas as células resultantes desse crescimento apresentem o mesmo genótipo, estabelecendo então uma nova linhagem celular.

Para separar as células uma a uma foi utilizada uma diluição em que continha aproximadamente 0.75 célula para cada 100 uL de meio. Utilizando uma pipeta multicanal, 100 uL de meio foi adicionado a placas de 96 poços previamente tratadas com gelatina 0.1%. Após uma semana, os poços foram analisados um a um por microscopia para averiguar a presença de colônias únicas. Os poços que possuíram colônias únicas foram passados para novas placas, tanto para o congelamento das células quanto para a extração de gDNA para testar se os clones eram positivos e posterior sequenciamento.

- **Extração de DNA genômico em placa**

Para extrair o DNA genômico dos clones, as células foram incubadas com 50 uL por poço de tampão de lise (10 mM Tris-Cl pH7.5, 10 mM EDTA pH 8.0, 10 mM NaCl, 0.5% Sarcosil e 1 mg/mL Proteinase K) durante 16h à 55°C. Foram adicionados 100 uL de uma solução de 75mM NaCl em etanol gelado e incubados por 30 minutos a temperatura ambiente. A placa então foi invertida em papel absorvente para a retirada das soluções e foi adicionado 100 uL de etanol 70% em cada poço. A placa foi invertida em papel absorvente e novamente adicionado 100 uL de etanol 70%. Depois a placa foi invertida mais uma vez e deixada para secar por alguns minutos a temperatura ambiente. Após a secagem o DNA precipitado foi ressuspensionado em 50 uL de água depc e incubado à 37°C durante pelo menos 1 hora. Para conservar, as placas contendo os gDNAs dos clones em cada poço foram guardadas em -20°C.

- **Diagnóstico dos clones**

- **Genotipagem por PCR seguida de sequenciamento (microdeleção)**

Após a extração do gDNA foi realizado o teste dos clones. Para os clones que sofreram microdeleções de exons, um PCR que visa amplificar a região abrangendo

a deleção foi realizado e os clones que apresentaram bandas indicando a ausência dos exons foram sequenciados.

- **Genotipagem por ensaio com a enzima T7 exonuclease seguida de sequenciamento (*frameshift*)**

Pelo fato dos clones que sofreram *frameshift* não possuírem deleções que permita a detecção por PCR, a região de interesse foi amplificada por PCR e reanelado 1:1 (150 ng : 150 ng) com o mesmo fragmento amplificado a partir de células selvagens, e submetidos a uma técnica que utiliza uma exonuclease T7 (New England Biolabs), capaz de cortar o DNA se houve um pareamento incompleto entre os fragmentos selvagens e dos clones. Com isso é possível a visualização de três bandas no gel para os clones que sofreram mutações.

4. RESULTADOS E DISCUSSÃO

1) Modulação da expressão de Coup-TFII em linhagens de CTE por silenciamento e superexpressão a partir de vetores virais de expressão gênica.

Para modular a expressão de COUP-TFII, obtivemos do laboratório do Dr. Okano⁶¹ os vetores de silenciamento shRNA COUP-TFII e de shRNA controle (*scramble*) que também possuíam como gene repórter a proteína fluorescente verde GFP. Também obtivemos o vetor que expressava COUP-TFII ectopicamente para super expressão do gene, mas este vetor não possuía um gene repórter, somente um gene que confere resistência à higromicina para as células transduzidas. A expressão de GFP pelas células transduzidas com shRNA COUP-TFII e shRNA controle permitiu a separação dessas células por FACS (do inglês Fluorescent Activated Cell Sorting), em citometria de fluxo. Após as transduções virais, as células TG2A foram passadas três vezes pelo processo de separação celular para garantir um alto grau de pureza das células transduzidas para os demais experimentos.

Na transdução viral por lentivirus, o vetor é incorporado ao genoma celular, podendo acontecer a integração de mais de um vetor por célula se essa for transduzida com mais de um vírus. Sendo assim, diferentes células poderiam ter números de cópias diferentes do vetor. Além de poder acarretar em variações não controláveis nos experimentos, é sabido que altas concentrações de GFP na célula pode causar efeitos tóxicos⁶⁵. Portanto, as células transduzidas foram separadas em duas populações de acordo com a quantidade de fluorescência que elas expressavam, a população 1 com menos fluorescência e a população 2 com mais fluorescência. A Figura 2 traz o padrão de fluorescência das células durante a primeira separação, portanto, antes que as células fossem separadas, após a transdução. É possível ver que apenas 1% das células haviam sido transduzidas com shCoup-TFII (Figura 2D), e 5% para o shControle (Figura 2C). As células GFP-positivas foram então separadas em população 1 (menos fluorescência) e população 2 (mais fluorescência).

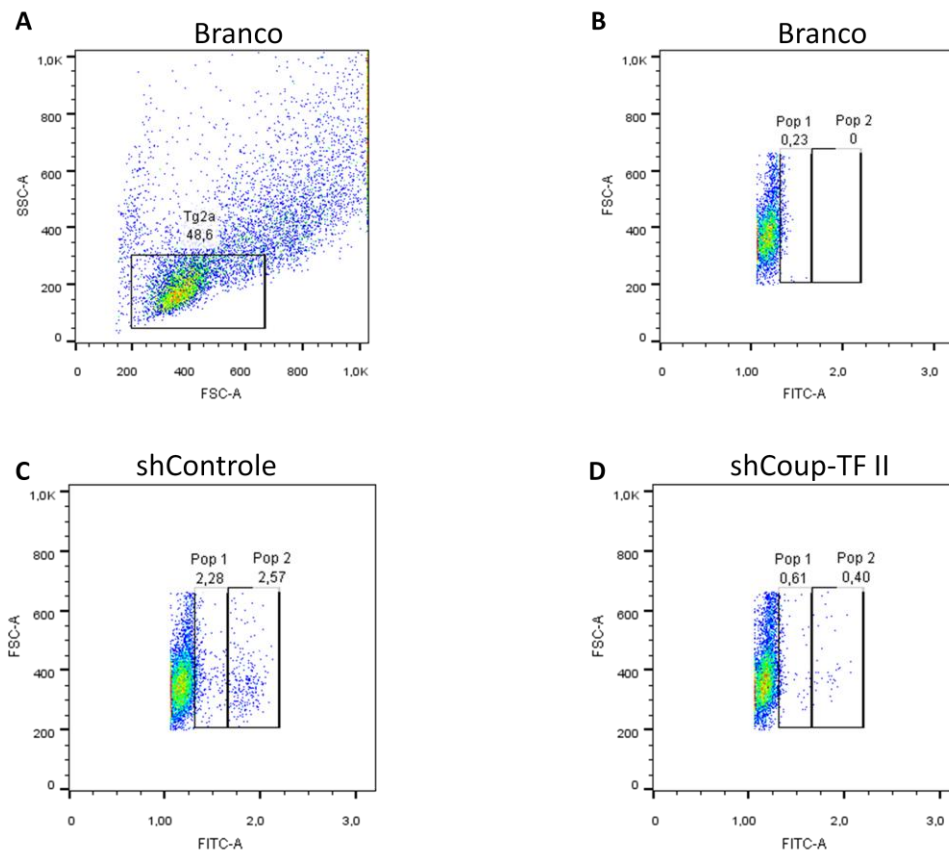


Figura 2: Primeira separação das células transduzidas. Gráficos gerados durante o FACS, reproduzindo como as células estavam antes da separação. Nos gráficos estão mostrados os limites para a seleção da população 1 (menos fluorescência) e a população 2 (mais fluorescência). A – Células brancas (não transduzidas), gráfico de tamanho das células (FSC-A) por textura das células (SSC-A), utilizado para selecionar o intervalo de células que serão analisadas. B – Células controle não transfectadas (“branco”), gráfico de fluorescência verde (FITC-A) por tamanho das células. C – Células shControle, gráfico de fluorescência verde por tamanho das células. D – Células shCoup-TFII, gráfico de fluorescência verde por tamanho das células.

Após a primeira separação, as células foram separadas em quatro grupos: shControle 1 (Figura 2C – Pop 1), shControle 2 (Figura 2C – Pop 2), shCoup-TFII 1 (Figura 2D – Pop 1) e shCoup-TFII 2 (Figura 2D – Pop 2). Estas células foram submetidas a uma segunda separação (Figura 4), que também revelou que em média 85% das células estavam mortas (Figura 3). Atribuímos este efeito à uma possível falha no processo de congelar e descongelar as células, resultando em um estresse acima do esperado. A Figura 3 mostra como as células estavam antes da segunda separação, evidenciando a porcentagem de células que estavam vivas e que estavam separadas em células únicas, podendo então ser separadas pelo FACS para a fluorescência.

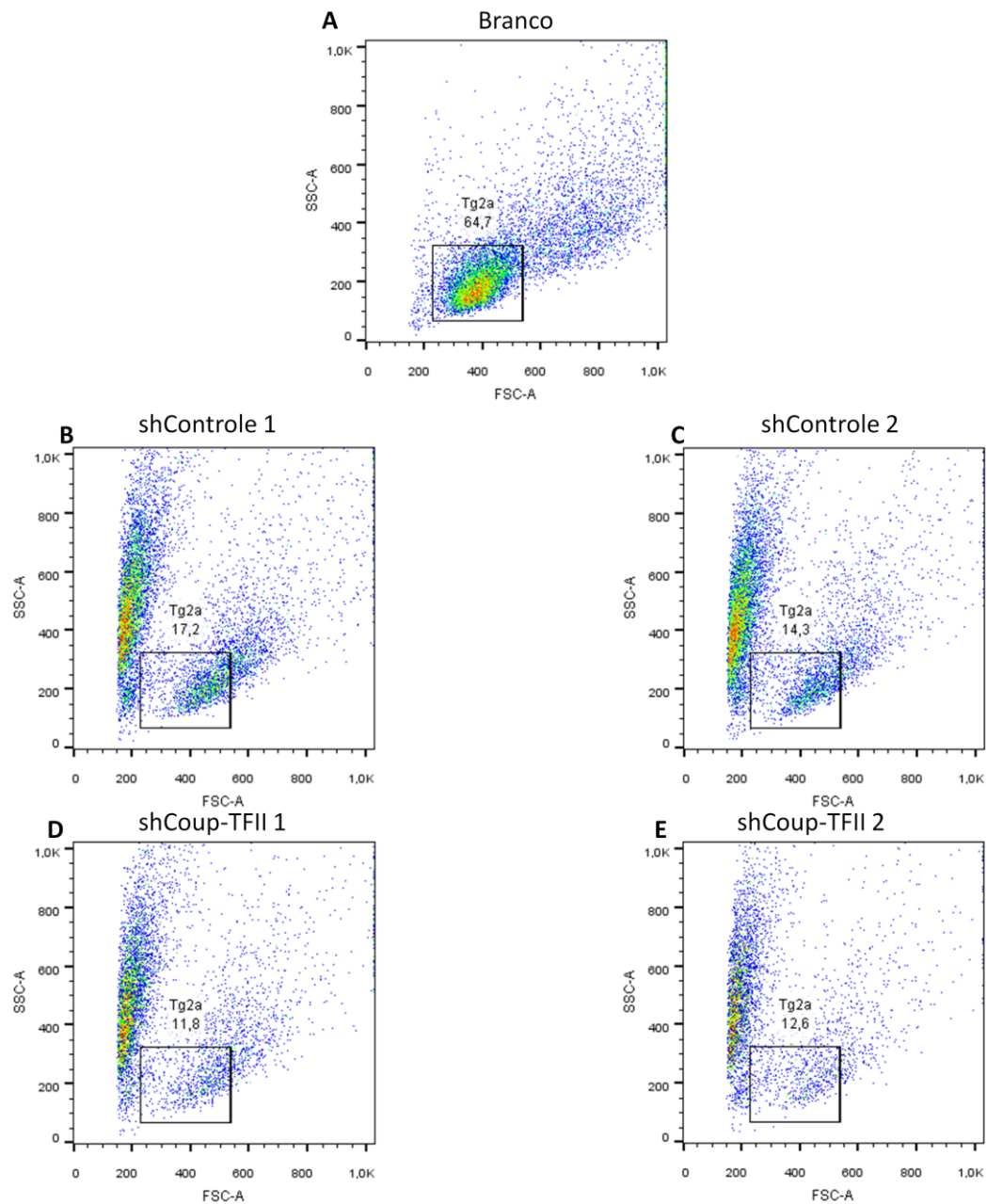


Figura 3: Gráficos de tamanho de células (FSC-A) pela granulosidade das células (SSC-A). A – Células controle não transduzidas, 64,7% B – Células shControle 1, 17,2%. C – Células shControle 2 14,3%. D – Células shCoup-TFII 1, 11,8%. E – Células shCoup-TFII 2, 12,6%.

Para a segunda e a terceira separação, as células shControle 1 (Figura 4A) foram selecionadas para separar somente as células dentro da população 1 (menos fluorescência), assim como as shCoup-TFII 1 (Figura 4C). E as células shControle 2 (Figura 4B) foram selecionadas para separar somente as células dentro da população 2 (mais fluorescência), assim como as shCoup-TFII 2 (Figura 4D). Sendo

assim, foi mantida a quantidade de expressão de GFP que foi separada nas populações desde a primeira separação.

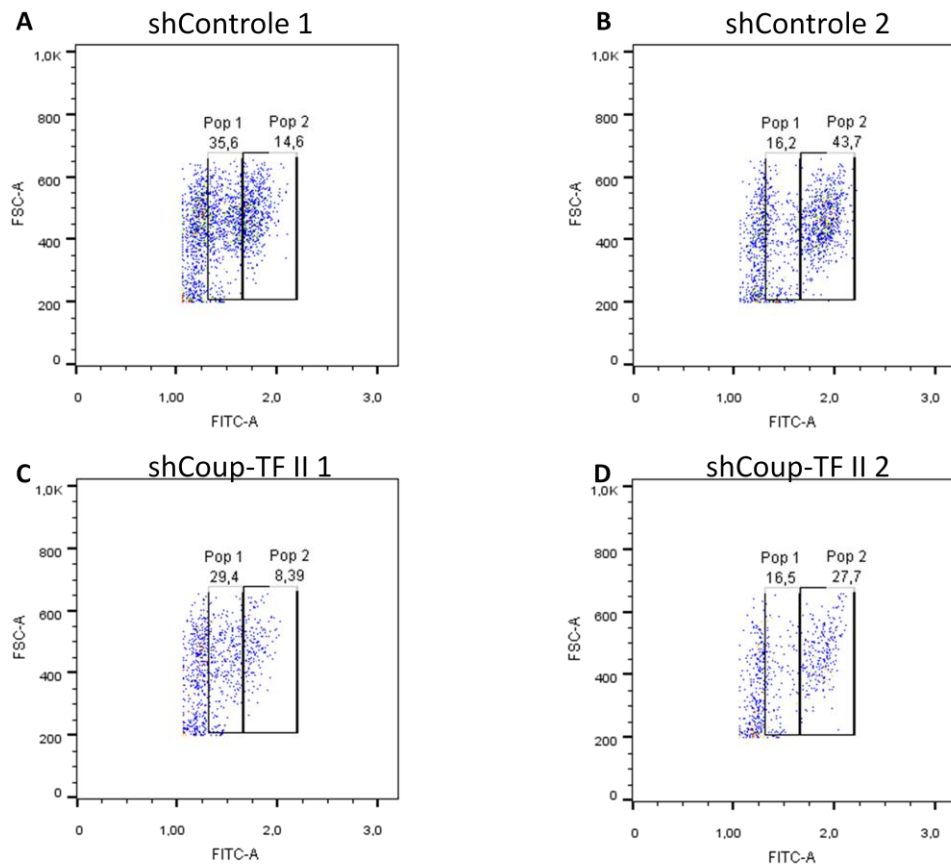


Figura 4: Segunda separação das células. Os gráficos mostrados foram feitos durante a separação, o que significa que representam a população das células antes da segunda. Os gráficos são de fluorescência verde (FITC-A) por tamanho das células (FSC-A). A – Células shControle 1. B – Células shControle 2. C – Células shCoup-TFII 1. D – Células shCoup-TFII 2.

Para quantificar a eficiência da separação, logo em seguida da segunda separação, as células shControle 2, que apresentavam a maior quantidade de células, passaram novamente pelo citômetro (Figura 5B), sendo possível comparar com a quantidade de células fluorescentes que haviam antes da segunda separação (Figura 5A) e posteriormente antes da terceira separação (Figura 5C).

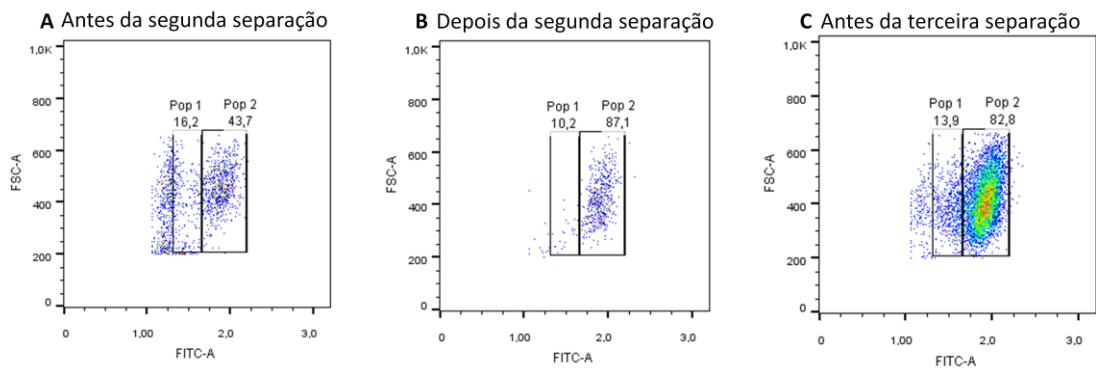


Figura 5: Células shControl 2 em gráficos de quantidade fluorescência verde (FITC-A) por tamanho das células (FSC-A), mostrando as porcentagens de células nas populações 1 e 2. A – Antes segunda separação, 43,7% de células na população 2. B – Depois da segunda separação, 87,1% de células na população 2. C – Antes da terceira separação 82,8% de células na população 2.

Apesar de separar as células duas vezes, não foi possível atingir níveis de expressão de GFP perto de 100% (Figura 5B e C), desejados para as análises de expressão gênica. Portanto, foi realizada uma terceira separação, seguindo os mesmos padrões das duas anteriores. A Figura 6 representa os gráficos gerados durante a terceira separação, quando 57,4% de células eram positivas para GFP na população 1 das células shControl 1 (Figura 6A); 82,8% das células eram positivas para GFP na população 2 das células shControl 2 (Figura 6B); e 84,5% das células eram positivas para GFP na população 2 das células shCoup-TFII 2 (Figura 6C). Assim como na segunda separação, as células shControl 1 separadas foram as células dentro da população 1, as células shControl 2 separadas foram as células dentro da população 2, e assim também para as células shCoup-TFII 1 e shCoup-TFII 2. A Figura 6 não mostra o gráfico shCoup-TFII 1, pois durante o FACS ocorreu um problema técnico, e o gráfico não foi salvo. Para evidenciar a diferença entre as populações 1 e 2 foi tirada uma imagem utilizando microscópio de fluorescência submetido à mesma exposição para a fluorescência verde, mostrando na Figura 7 que há uma diferença visual nítida na expressão dos vetores shRNA nas diferentes populações representada por GFP.

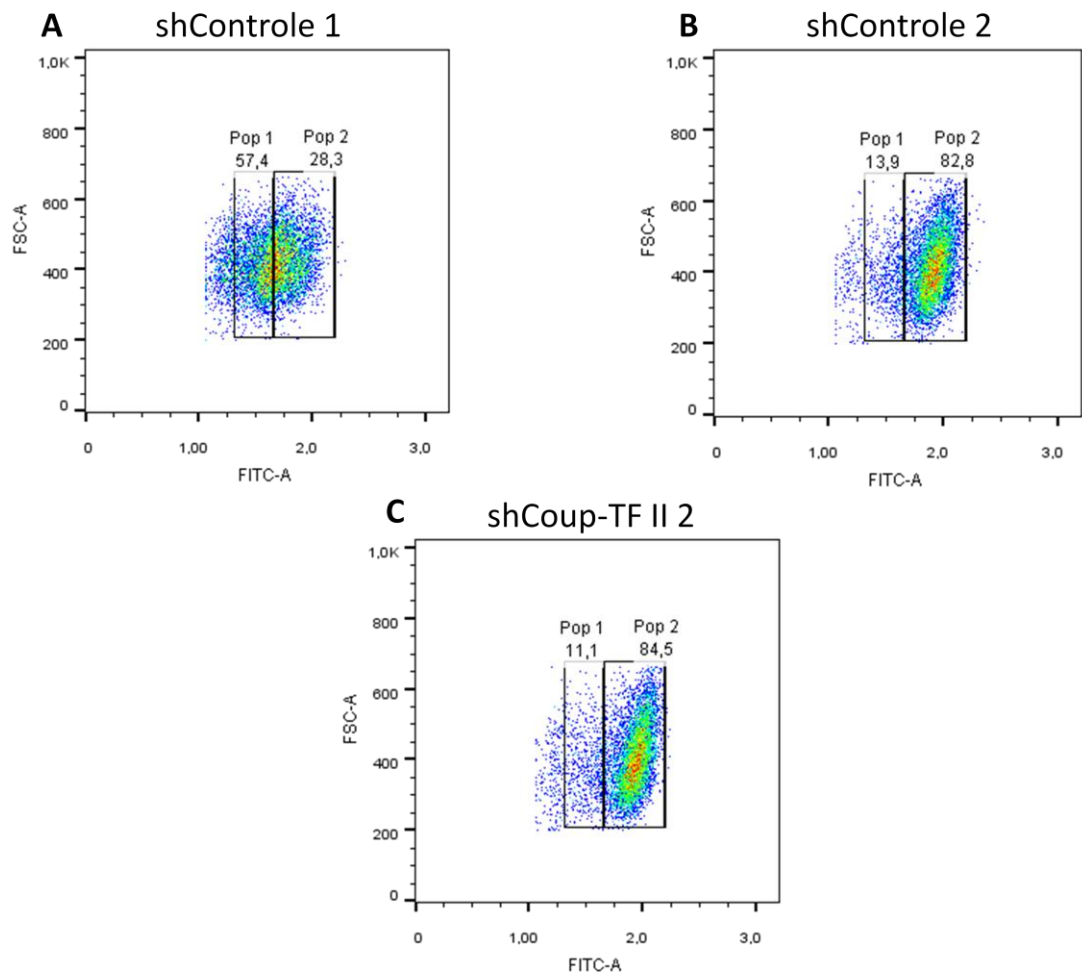


Figura 6: Terceira separação, gráficos de quantidade de fluorescência verde (FITC-A) por tamanho das células (FSC-A). A – Células shControle 1. B – Células shControle 2. C – Células shCoup-TFII 2.

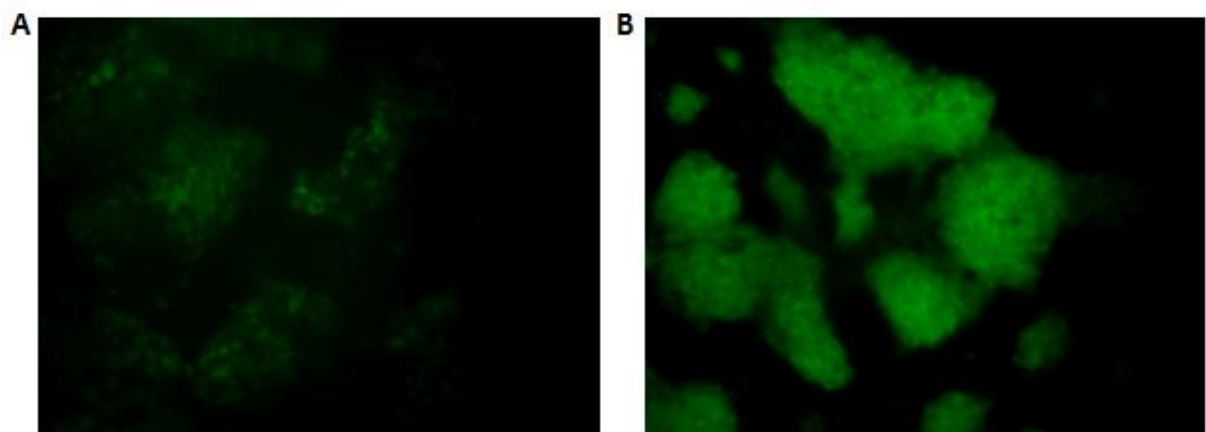


Figura 7: Microscopia de fluorescência evidenciando a diferença de expressão de GFP entre as populações shCoup-TFII 1 (A) e shCoup-TFII 2 (B) em imagens tiradas no mesmo tempo de exposição.

2) Ensaio de diferenciação a partir de CTE moduladas para a expressão de COUP-TFII.

Após a terceira separação, as células foram submetidas à diferenciação em camadas únicas (*monolayer*) para serem analisadas quanto à expressão de Coup-TFII modulando as expressões de outros genes. Para as células transduzidas com o vetor de superexpressão, foi adicionado o antibiótico higromicina e as células foram mantidas nesse antibiótico por várias passagens. Nesta etapa, foi realizado o procedimento denominado pré-plaqueamento, que toma vantagem do diferencial de capacidade de adesão entre células indiferenciadas e já em processo de diferenciação. Neste procedimento, as células dissociadas foram incubadas em placa gelatinizada por 30 minutos, tempo suficiente para que as células já em processo de diferenciação se aderissem na placa enquanto as células indiferenciadas permanecem em suspensão e são portanto coletadas com o meio de cultura.

As seis populações de células estudadas (tipo selvagem, shControle 1, shControle 2, shCoup-TFII 1, shCoup-TFII 2 e Superexpressão de Coup-TFII) foram diferenciadas até o sexto dia de diferenciação. As células foram então fixados para análise por imunohistoquímica e o RNA foi extraído para análise de expressão gênica por qPCR. Por imunohistoquímica (Figura 8) foi observado a coloração do núcleo por Dapi (azul), do GFP (verde) expresso pelas células transduzidas e da marcação por anticorpo contra COUP-TFII (vermelho). As populações shControle 2 (Figura 8A), shCoup-TFII 2 (Figura 8B) e superexpressando Coup-TFII (Figura 8D) foram analisadas por microscopia confocal de fluorescência (INFABIC) e revelaram que a expressão de Coup-TFII é co-localizada com a fluorescência verde nas células shControle (Figura 8A) mostrando assim, que não há influências na expressão do gene quando o sistema sh é utilizado nas células. No entanto, também foram observados células expressando a fluorescência verde (por tanto deveriam estar silenciadas para Coup-TFII) com expressão aparentemente normal para Coup-TFII (Figura 8B).

Além disso, esperava-se que transdução das células com o vetor de superexpressão de Coup-TFII iria resultar em (1) mais células expressando Coup-

TFII do que em populações do tipo selvagem e (2) maiores intensidades de expressão de Coup-TFII nestas células. Entretanto, nenhum dos dois efeitos foram observados em nosso experimento de superexpressão de Coup-TFII (Figura 8D).

Por conta dos resultados obtidos com a imunohistoquímica, as análises por qPCR não foram realizadas.

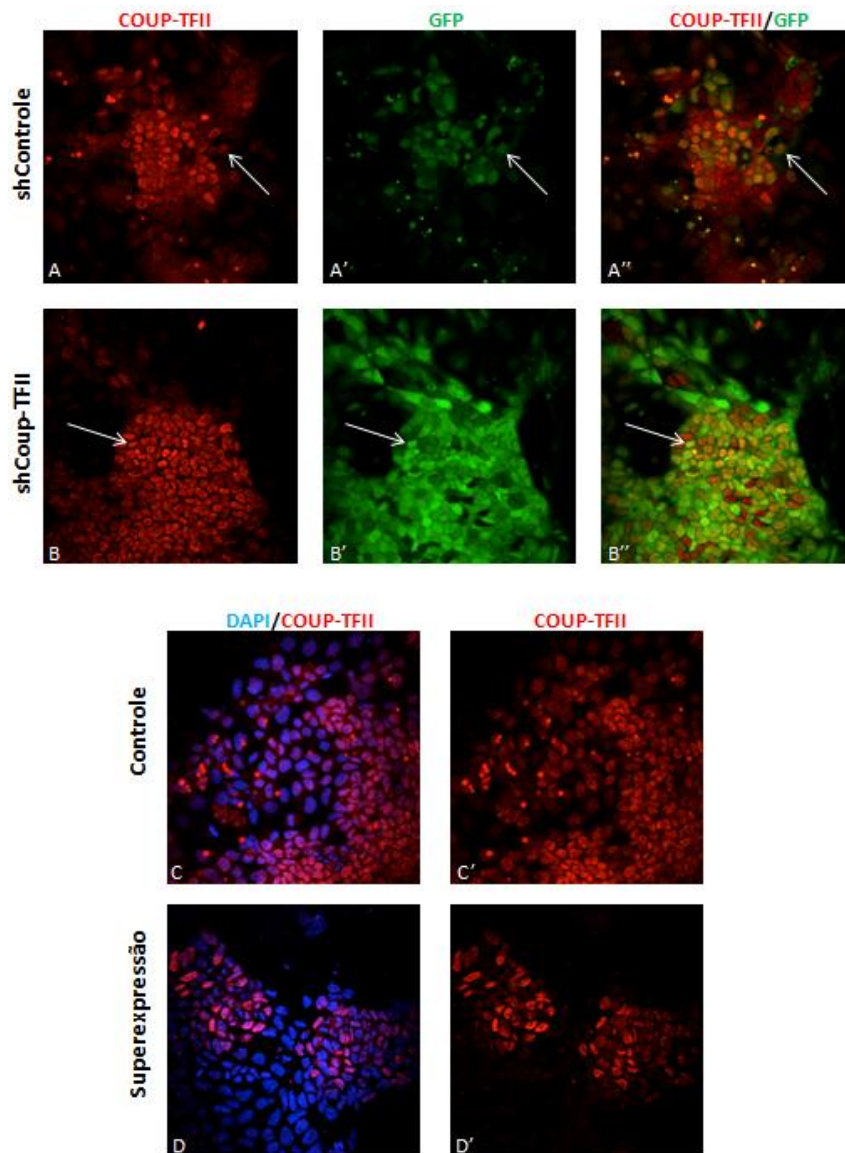


Figura 8: Imunohistoquímica para COUP-TFII (vermelho) em células transduzidas com vetores de shControl (GFP – Verde), shCoup-TFII (GFP – Verde) e superexpressão Coup-TFII diferenciadas por 5 dias. O núcleo foi evidenciado por DAPI. A- Células shControl. B – Células shCoup-TFII. C – Célular Controle. D – Células transduzidas com vetor de superexpressão de Coup-TFII.

3) Estabelecimento de CRISPR/Cas9 e geração de linhagens CTE nocaute para COUP-TFII.

a) Padronização da lipofecção

Para introduzir os vetores dos CRISPR/Cas9n (ver material e métodos) nas células TG2A primeiramente foram testadas diversas condições de lipofecção para determinar qual condição seria a mais adequada. Nove combinações de Lipofectamina 2000 com um vetor controle de expressão de GFP foram utilizadas, variando as quantidades de Lipofectamina 2000 em 0.25uL (Figura 9A, D e G), 0.625uL (Figura 9B, E e H) e 1.25uL (Figura 9C, F e I); e as concentrações de vetor em 100ng (Figura 9A-C), 312ng (Figura 9D-F) e 625ng (Figura 9G-I). Os testes foram feitos em células plaqueadas em poços de placas de 96 poços (80 mil células por poço). Um dia após as lipofecções, as células foram analisadas por citometria de fluxo (Figura 9).

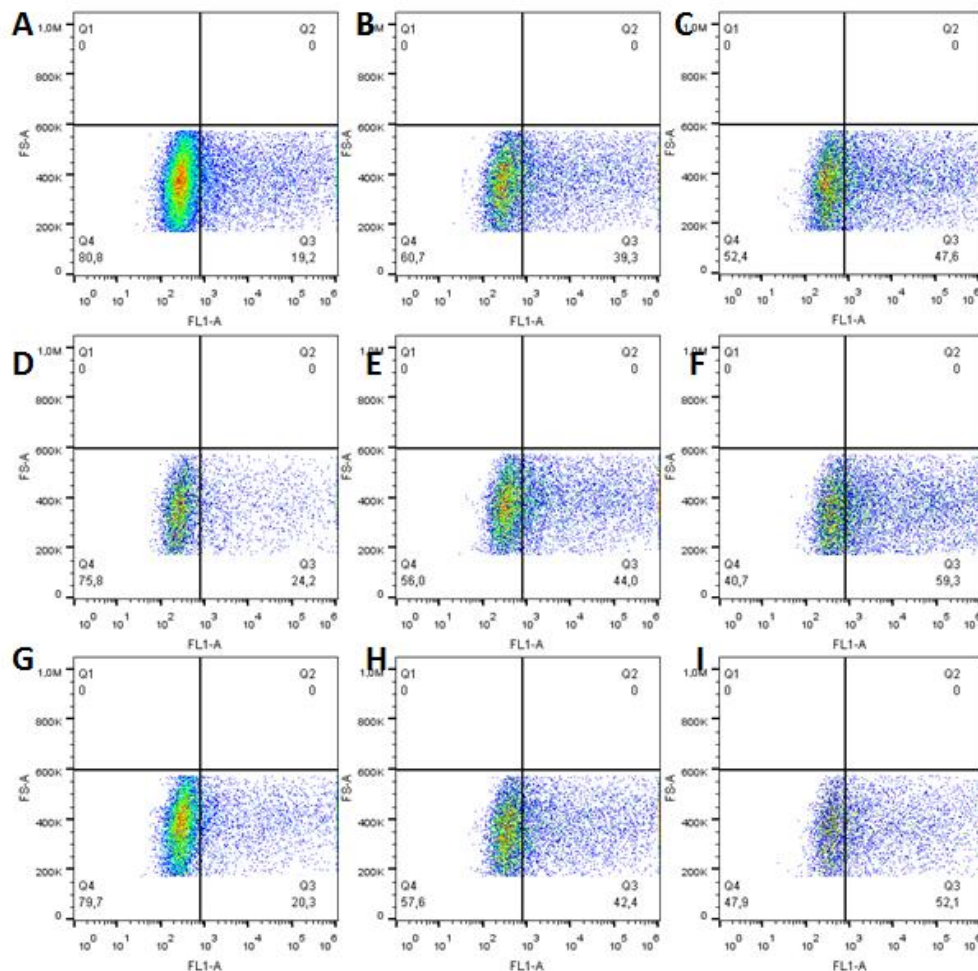


Figura 9: Citometria das diferentes condições testadas para lipofecção de células da linhagem E14Tg2a. A – 0.25uL de Lipofectamina e 100ng de vetor (19.2% de células GFP); B - 0.625uL de

Lipofectamina e 100ng de vetor (39.3% de células GFP); C - 1.25uL de Lipofectamina e 100ng de vetor (47.6% de células GFP); D - 0.25uL de Lipofectamina e 312ng de vetor (24.2% de células GFP); E - 0.625uL de Lipofectamina e 312ng de vetor (44% de células GFP); F - 1.25uL de Lipofectamina e 312ng de vetor (59.3% de células GFP); G - 0.25uL de Lipofectamina e 625ng de vetor (20.3% de células GFP); H - 0.625uL de Lipofectamina e 625ng de vetor (42.4% de células GFP); I - 1.25uL de Lipofectamina e 625ng de vetor (52.1% de células GFP).

A partir dos resultados foi possível observar que entre 100ng de vetor e 625ng de vetor a média de variação é de 2.9% enquanto que entre 0.25uL de lipofectamina e 1.25uL de lipofectamina a média da variação é de 31.7%. A partir desses resultados foi possível concluir que a maior influência na qualidade da Lipofecção é a quantidade de lipofectamina e não a quantidade de vetor. A partir desta observação as quantidades de lipofectamina e vetor foram padronizadas para a lipofecção das células para CRISPR/Cas9.

b) Diagnóstico dos clones

i. Microdeleção do exon 2 do COUP-TFII

A estratégia para a microdeleção do exon 2 de Coup-TFII envolve a lipofecção das TG2A com os sgRNA 15, 16, 17 e 18 (Tabela 1). Esses guias estão localizados em regiões flanqueando o exon 2 e a ação da enzima Cas9n resultará na completa deleção deste exon (Figura 11). Como o exon 2 possui aproximadamente 500pb, o diagnóstico dos clones consistiu na amplificação por PCR onde um primer Forward (102 – Tabela 1) se anela no intron 1 e um primer Reverse (103 – Tabela 1) se anela no intron 2. Em clones tipo selvagem era esperado um amplicon de aproximadamente 1700 pb, enquanto que para clones onde ocorreu a microdeleção era esperado uma banda de aproximadamente 1000 pb (Figura 10). Como controle da reação foi utilizado primers para a amplificação de transcritos do gene da Beta-Actina (Tabela 1). Dos 26 clones testados, foram identificados 13 clones com microdeleções em homozigose (exemplos em Figura 10, amostras #2, #6, #18 e #20) e 1 em heterozigose (exemplo em Figura 10, amostra #16). Após a comprovação das amplificações por eletroforese em gel de agarose, as amostras foram submetidas ao sequenciamento, evidenciando a perda do exon 2 e de uma parte da região intrônica 5' e 3' ao exon 2 (Figura 11).

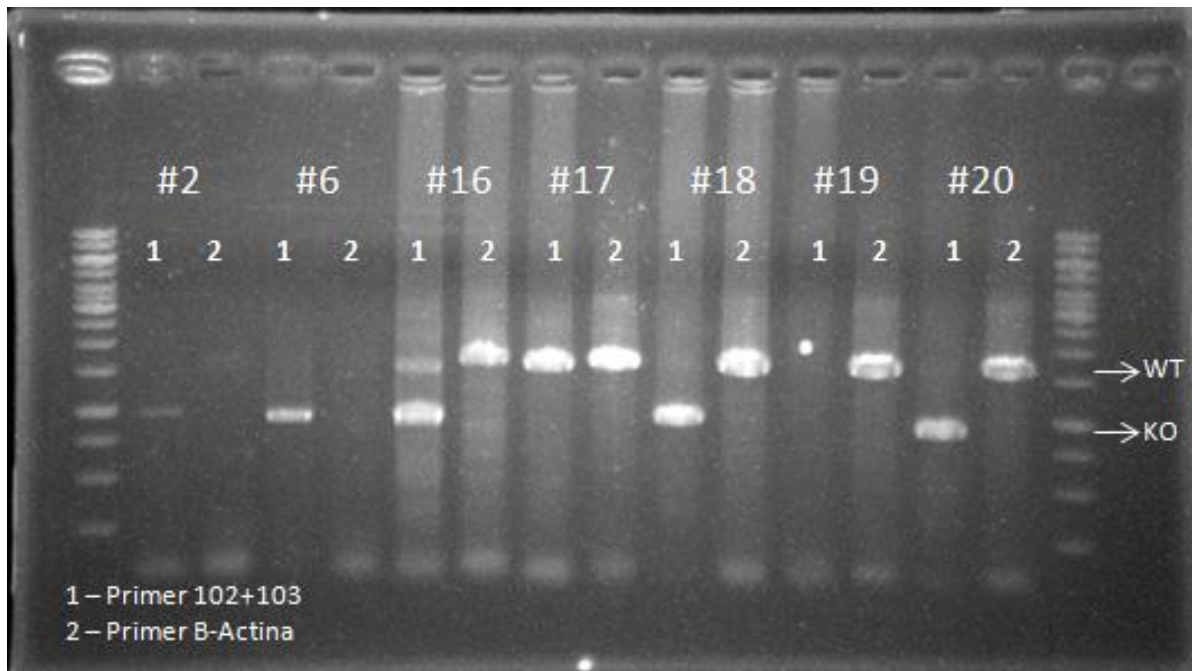


Figura 10: Gel de PCR para diagnóstico dos clones que sofreram microdeleção do exon 2 do COUP-TFII. Os PCRs foram feitos utilizando primers para diagnóstico e primers para B-Actina (controle positivo do PCR). Os tamanhos esperados para os clones que sofreram a microdeleção é de aproximadamente 1000 pb e para os selvagens é de 1700 pb.

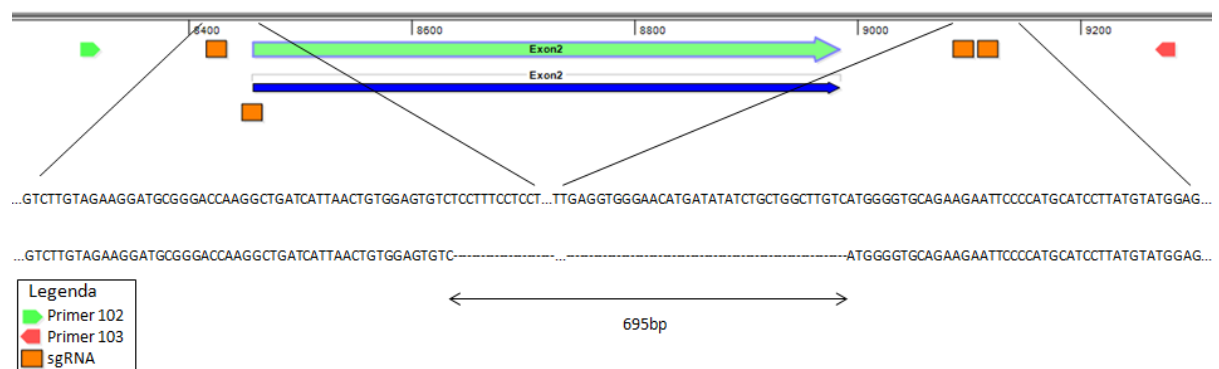


Figura 11: Resultado do sequenciamento de um dos clones quando comparado à sequência tipo selvagem do exon 2 do COUP-TFII. A figura também mostra os primers utilizados para diagnóstico e os sgRNAs

ii. Microdeleção dos exons 1 e 2 do TV1 do COUP-TFII

Esta estratégia de microdeleção envolveu a utilização dos sgRNAs 03, 04, 17 e 18. Estes guias estão localizados na região 5' UTR (sgRNAs 03 e 04) e no segundo intron (sgRNAs 17 e 18), ao passo que a ação da enzima Cas9n resultou na deleção simultânea dos exons 1 (específico para TV1) e 2. Após o isolamento dos clones, os mesmos foram genotipados para uma região que se estende desde o

primeiro exon (região UTR 5') (Primer 98 – Tabela 1) até depois do segundo exon (Primer 107 – Tabela 1) (Figura 12). Para os clones do tipo selvagem, que não sofreram mutação, era esperado um amplicon de aproximadamente 3600pb. Para os clones que sofreram a microdeleção dos dois exons era esperado um amplicon de aproximadamente 900pb.

Dos 22 clones testados, 5 apresentaram bandas na altura esperada para nocautes (Figura 12, #50, #51, #55, #56 e #63). Os demais clones apresentaram bandas de tamanhos maiores, menores, três bandas ou nenhuma banda. Os cinco clones com bandas na altura esperada para a deleção e os clones #21 e #67 (Figura 12), que apresentavam bandas únicas e menores do que o esperado para as microdeleções foram sequenciados.

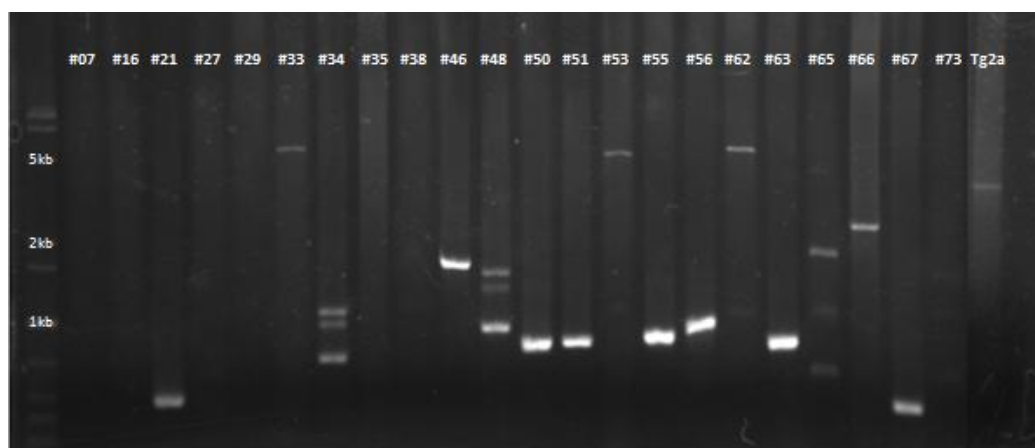


Figura 12: Gel de agarose correspondente ao PCR para diagnosticar a microdeleção em clones que foram transfectados com os sgRNAs para deletar os exons 1 e 2 da TV1.

O sequenciamento dos clones comprovou que houve uma deleção de aproximadamente 2733pb nos clones nocautes #50, #51, #55 e #56 (Figura 13), enquanto que os clones #21 e #67 tiveram aproximadamente 3080pb deletados.

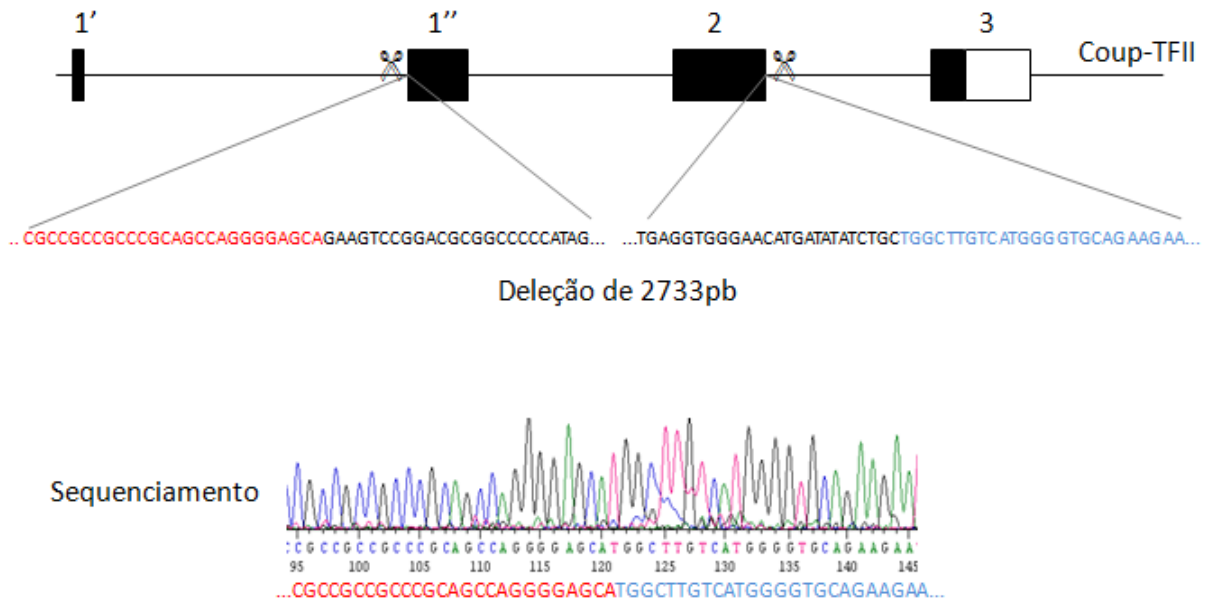


Figura 13: Sequenciamento do clone #50 representando os resultados dos sequenciamentos dos clones #51, #55 e #56 para a microdeleção dos exons 1 e 2.

iii. *Frameshift* a partir do exon 1 do COUP-TFII TV1.

Esta estratégia teve como objetivo causar a quebra da dupla fita do DNA próximo ao códon de iniciação do COUP-TFII, o que resultou em deleções e inserções de nucleotídeos pela maquinaria celular enquanto esta tentava reparar a quebra. Como resultado dessas inserções e deleções o gene poderia alterar o código de leitura, e a proteína resultante não seria funcional. Os alelos gerados pelo método de alteração de código de leitura (*Frameshift*) não permitem diagnósticos por amplificação, uma vez que esperavam-se deleções de poucos nucleotídeos, o suficiente para alterar o código de leitura de tradução. Portanto, foi utilizado o método de diagnóstico com a enzima T7 endonuclease, onde as fitas de DNA resultantes da amplificação do local do *Frameshift* foram reaneladas com fitas de DNA tipo selvagem referente à mesma região de amplificação. As fitas resultantes desse reanelamento tiveram todos os nucleotídeos pareados entre si, caso o clone testado fosse do tipo selvagem. Porém, se o clone testado apresentasse pelo menos um nucleotídeo diferente do tipo selvagem, as fitas não iriam se complementar perfeitamente, e a enzima T7 endonuclease iria reconhecer este pareamento imperfeito no DNA e clivar este sítio. Desta forma, quando o DNA pareado com o tipo selvagem apresentava alguma mutação e sofria a clivagem, o fragmento

proveniente da amplificação do PCR clivado apresentava bandas extras, referentes aos fragmentos menores clivados, e então era possível observar por gel de eletroforese bandas abaixo da banda referente ao alelo tipo selvagem, o que indicava a existência de pelo menos um alelo mutado neste clone. A Figura 14 representa o teste de três clones, e foi mostrado como as bandas estavam antes da adição da enzima T7 (-) e depois da adição da enzima T7 (+). A Partir dessa figura foi possível observar que existia banda extra no clone 0304_31, enquanto essa banda não era presente nos outros clones.

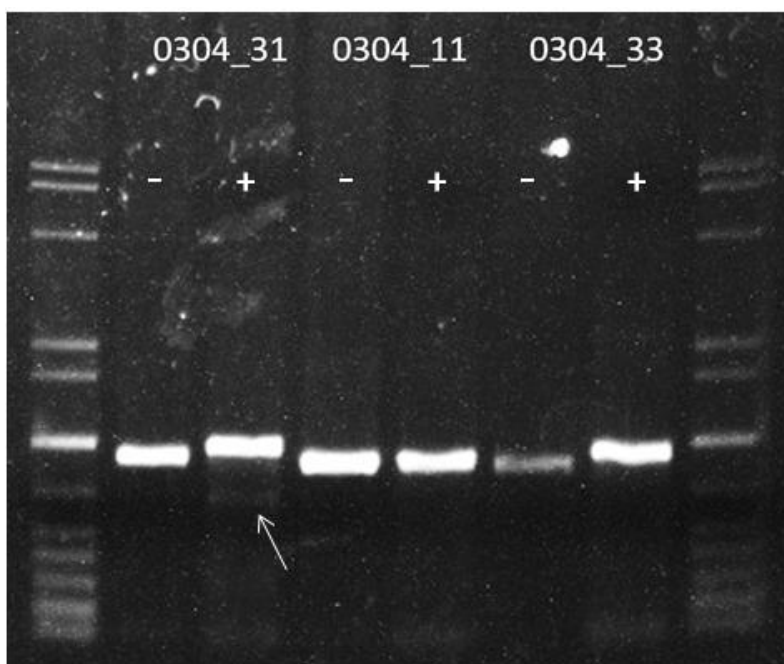


Figura 14: Gel de agarose evidenciando a presença de bandas extras após o tratamento com a enzima T7.

Esse clone foi sequenciado a partir do PCR e a Figura 15 mostra o resultado. O sequenciamento foi feito tanto no sentido forward utilizando o primer 92 (Tabela 1) quanto no sentido reverse utilizando o primer 93 (Tabela 1). Foi possível observar que nos dois sentidos os picos do sequenciamento estavam bons e claros até o momento em que encontram a região do sgRNA4 (Figura 15). A partir deste momento (região destacada em amarelo na Figura 15) os picos se encontravam confusos e misturados. Este erro de leitura está atribuído ao fato de termos sequenciado o fragmento purificado da PCR, resultante de uma mistura dos alelos que estavam ali presentes. Neste caso, a presença de alelos com deleções diferentes na reação resultaram em leituras múltiplas pelo sequenciamento. Como

as regiões flanqueando a região onde se esperava a deleção apresentavam um bom sequenciamento, concluímos que houve alguma alteração no genoma no local desejado. No momento seria necessário clonar o produto de PCR em vetores e transformando em bactérias para que o sequenciamento seja realizado a partir de uma única sequência molde, ou alelo.

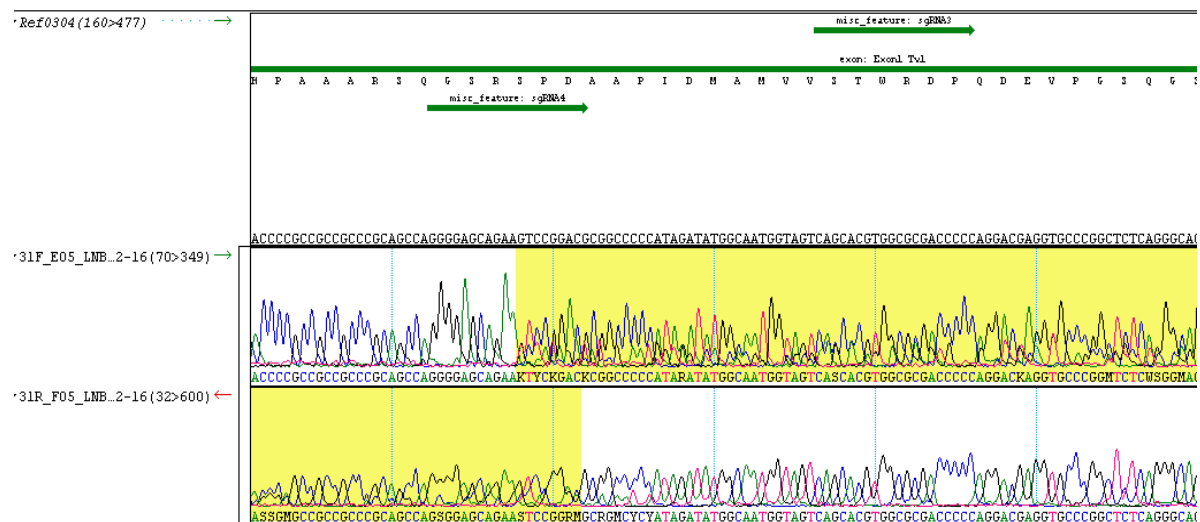


Figura 15: Resultado do sequenciamento do clone 31 a partir do produto do PCR utilizando primer forward (92 – Tabela 1) e reverse (93 – Tabela 1).

4) Ensaio de diferenciação celular com CTE nocaute para COUP-TFII

Diferenciações em CE foram realizadas utilizando as células selvagens e células do clone nocaute denominado 1/33, que sofreu a microdeleção do exon 2 (Figura 11). Após seis dias de diferenciação em CE, uma imunohistoquímica foi realizada utilizando anticorpo contra a proteína do Coup-TFII e não foi possível detectar a presença da proteína nos CE nocautes 1/33, como mostra a Figura 16.

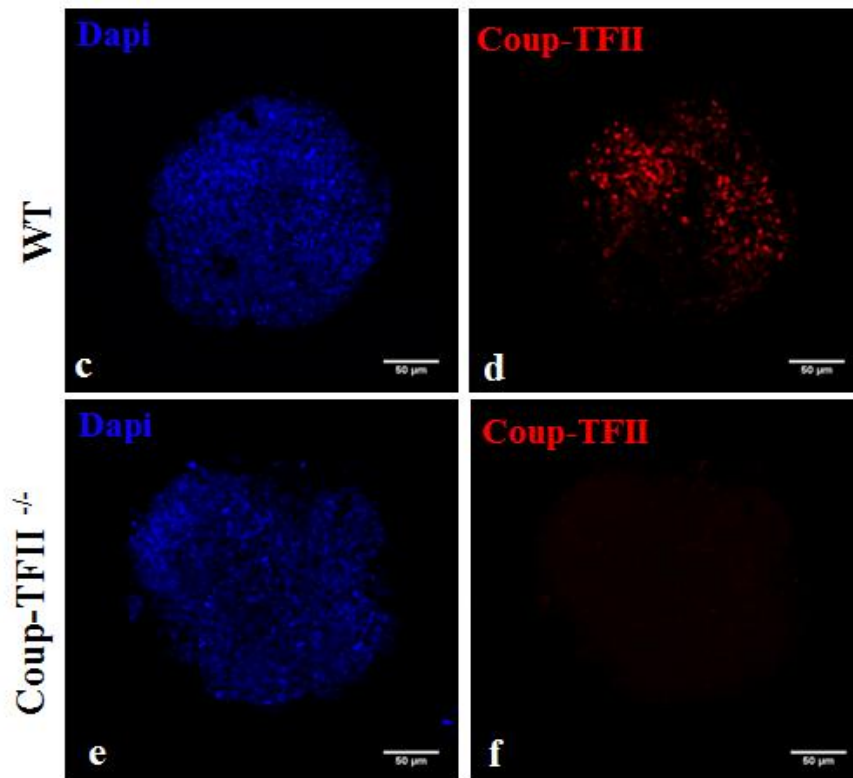
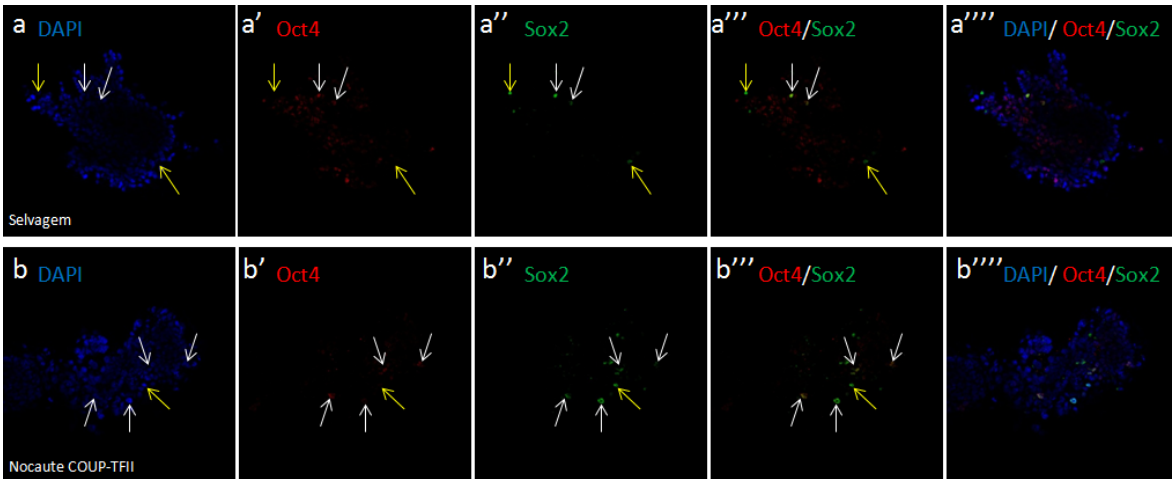


Figura 16: Imunohistoquímica em CEs D6, utilizando anticorpo contra a proteína do Coup-TFII.

Os resultados apontam para um possível aumento da expressão de Sox2 e da co-expressão de Sox2 e Oct4 nas células nocautes para Coup-TFII (Figura 17). Sabe-se que células que expressam tanto Oct4 quanto Sox2 são consideradas pluripotentes^{2, 66}, enquanto que células que mantêm somente a expressão de Sox2 são direcionadas para ectoderme⁶⁶. Estes resultados sugerem que o nocaute de COUP-TFII atrasa a saída das células da pluripotência. Enquanto CEs diferenciados por 4 dias geralmente apresentam cerca de 24% de células marcadas para Sox2 também positivas para Oct4 (Figura 17 A e C), nos CEs formados a partir das células mutadas para Coup-TFII cerca de 50% das células marcadas para Sox2 são também duplo-positivas para Oct4 (Figura 17 B e C).



c

Linhagem	Sox2	Sox2/Oct4	% Sox2/Oct4
Nocaute	340	169	50
Selvagem	259	63	24

Figura 17: Imunohistoquímica em CEs diferenciados por 4 dias. Setas brancas apontam células que co-expressam Sox2 e Oct4, e setas amarelas apontam células que expressam somente Sox2. A – CEs de células selvagens. B - CEs de células nocautes. C – Gráfico mostrando a quantidade de células que expressam Sox2 e células que expressam Sox2 com Oct4.

Visando estudar a expressão diferencial de genes marcadores de pluripotência e tipos celulares específicos, coletamos amostras de RNA de três ensaios de diferenciação independentes, para os clones 1/1 e 1/33 (microdeleção do exon 2) e realizamos qPCR em triplicatas técnicas de cada uma das replicatas biológicas (Figura 18). Observamos uma tendência de elevação da expressão de Oct4 que, apesar de não ser estatisticamente significativa, corrobora com resultados já publicados⁴². Portanto, estes resultados apontam para um sucesso no nocaute de Coup-TFII através de um fenótipo diferencial quando comparado ao tipo selvagem.

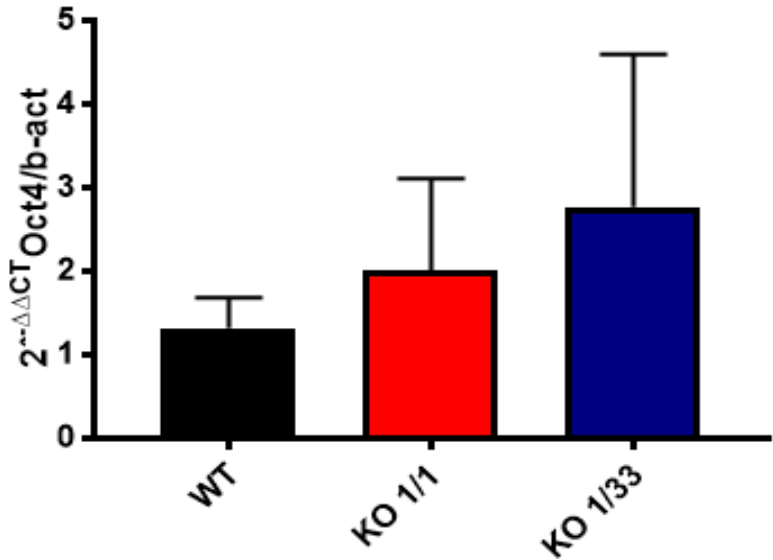


Figura 18: qPCR para o gene Oct4 utilizando B-actina como controle endógeno da reação, três linhagens de células selvagens (WT), e duas de nocautes (1/1 e 1/33). As diferenciações foram feitas por 6 dias em CE, e repetidas três vezes para a geração desse gráfico.

5) Estabelecimento de uma linhagem repórter COUP-TFII-mCherry por CRISPR/Cas9

A estratégia para a geração de uma linhagem repórter através de CRISPR/Cas9 envolveu o corte da Cas9n próximo ao códon de parada do gene e que o vetor doador de homologia, isso é, um vetor que continha regiões homólogas (braço homólogo 5' e braço homólogo 3') às regiões flanqueando o local de inserção do gene repórter (mCherry), para que o gene repórter fosse inserido imediatamente após o códon de parada (Figura 19).

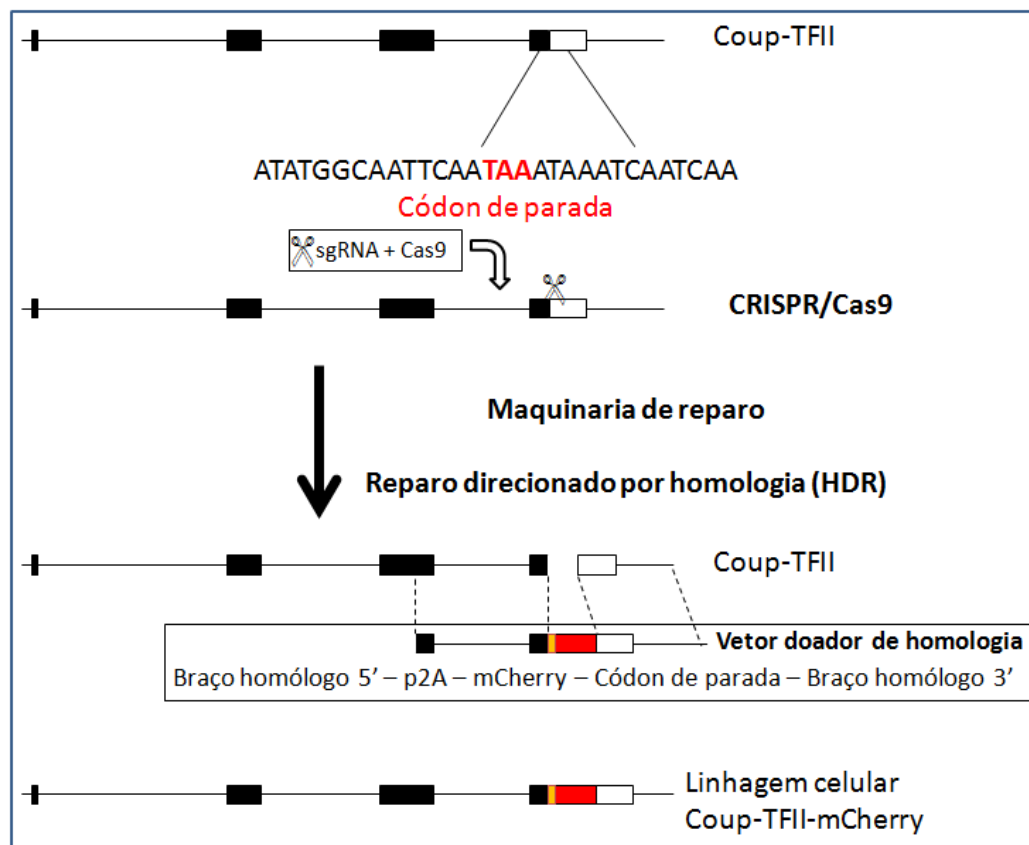


Figura 19: Esquema representando a estratégia e os passos para o estabelecimento da linhagem repórter COUP-TFII-mCherry

a) Construção do vetor doador de homologia.

As sequências homólogas flanqueando o local de inserção do gene mCherry foram clonadas a partir do DNA genômico extraído das próprias células E14 Tg2a, onde os primers para amplificar cada braço homólogo foram desenhados de forma que, após a amplificação, os braços contivessem enzimas de restrição que permitam a clonagem destas sequências em um vetor e a posterior clonagem em ordem, dos fragmentos do braço 5' (Figura 20A), mCherry (Figura 20B) e braço 3' (Figura 20C) (Primers 21, 23, 24, 25, 34 e 35 – Tabela 1). Cada um dos três fragmentos (braços 3', 5' e mCherry) foram clonados separadamente no vetor pBluescript e transformados em bactérias para que, a partir de uma única colônia, pudesse ocorrer o sequenciamento dos mesmos individualmente. As sequências dos primers utilizados está na Tabela 1 e um esquema contendo as enzimas de restrição está na Figura 20. O mCherry foi clonado a partir do vetor 26031:543-p3E-2A-mcherryA (Addgene). Junto à sequência do mCherry havia uma sequência que codificava para o peptídeo 2A, que tem a capacidade de se auto-clivar. Com isso, a tradução da proteína do COUP-TFII e do mCherry seria sequencialmente, mas as duas proteínas se tornariam independentes após a clivagem do peptídeo 2A.

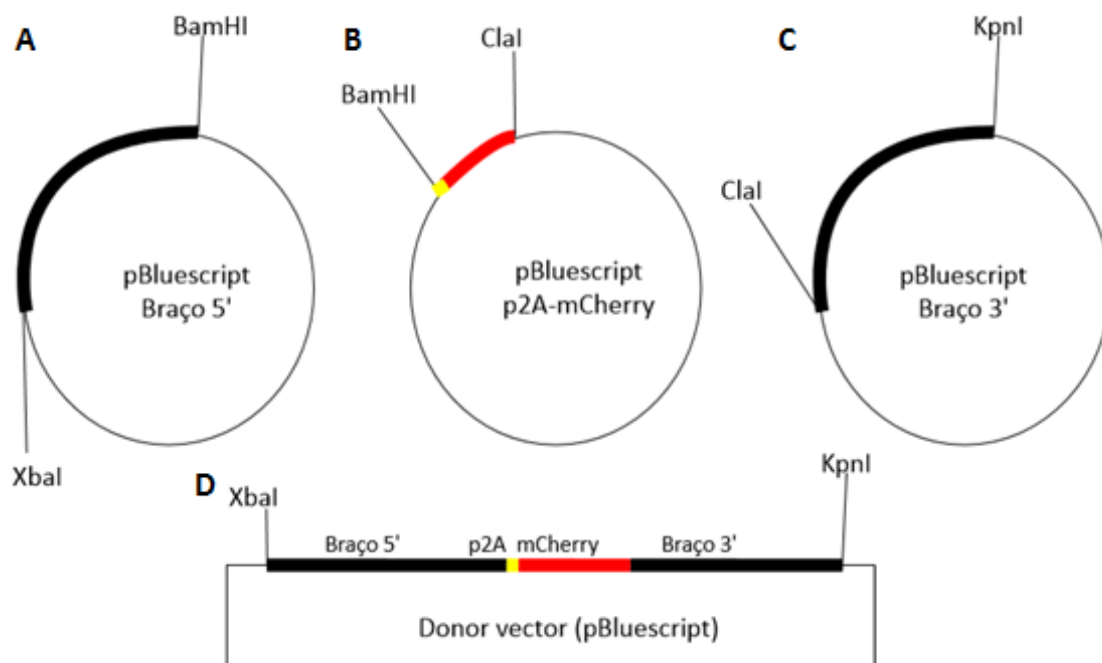


Figura 20: Esquema mostrando os braços 5', 3' e mCherry clonados individualmente no vetor pBluescript e depois todos em sequência montando o donor vector (vetor doador de homologia)

Após o PCR para amplificar os braços homólogos (Figura 21) e o gene mCherry (Figura 22), os amplicons foram purificados e digeridos com enzimas de

restrição de acordo com a Figura 20A-C. Os fragmentos foram então ligados separadamente ao pBluescript, que também havia sido digerido com as mesmas enzimas. Após o sequenciamento dos vetores, digestões com as enzimas como mostradas na Figura 20A-C ocorreram e para que então os fragmentos fossem ligados formando um único vetor doador de homologia contendo os três fragmentos. Porém, após diversas tentativas de fazer a ligação dos três fragmentos sem sucesso optamos por enviar a sequência completa, mostrada na Figura 20D, para ser sintetizada pela empresa GenScript.

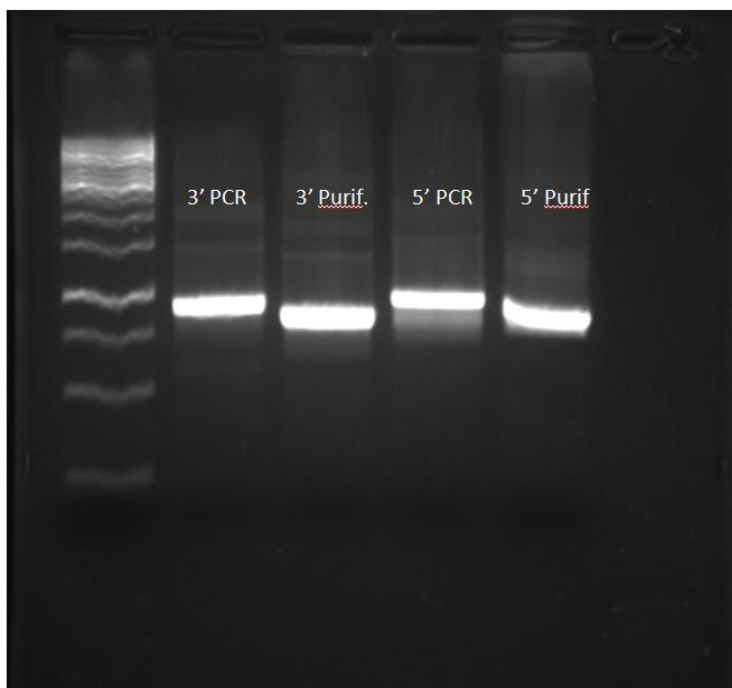


Figura 21: Gel de agarose mostrando os braços 3' e 5' após o PCR e purificação de cada um para a clonagem no vetor pBluescript. Ambos os braços apresentam em torno de 1000 pb

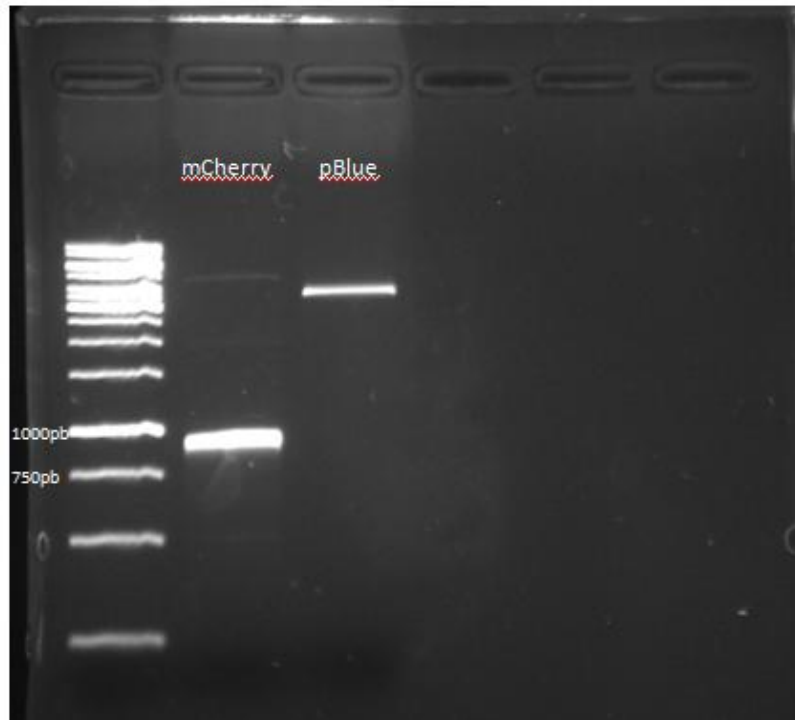


Figura 22: Gel de agarose mostrando o mCherry após o PCR (altura esperada de 800pb) e o vetor pBluescript linearizado (altura esperada de 3000 pb).

b) Geração das linhagens e diagnóstico dos clones

Uma vez recebido o vetor doador de homologia, as células foram lipofectadas com este vetor e com os sgRNAs (sgRNA1 e 11 – Tabela 1), os clones selecionados foram genotipados por PCR para a inserção do mCherry, amplificando-se a região em torno de onde o mCherry deveria ter sido incluído (Primer 108 e 111 – Tabela 1). Esperava-se que o tipo selvagem apresentasse aproximadamente 300pb e que os clones que tiveram o mCherry inserido apresentasse aproximadamente 1100pb.

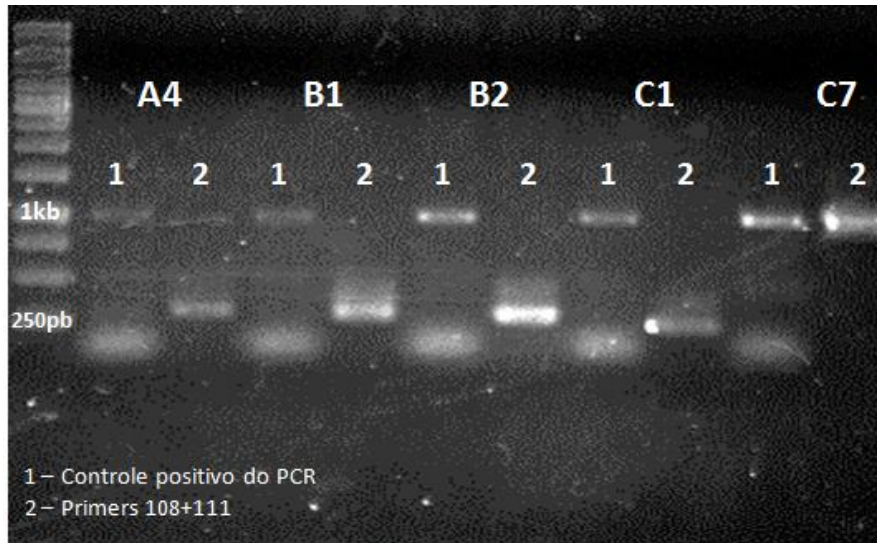


Figura 23: Gel de agarose do PCR com os primers 108 e 111 para diagnosticar os clones que tiveram a inserção do 2a-mCherry no genoma. O PCR foi feito utilizando (1) os primers para amplificar o braço homólogo 3' como controle positivo e (2) os primers para amplificar a região flanqueando o local de inserção do mCherry. Para os primers 2, os clones que apresentam aproximadamente 300pb não tiveram a inserção do mCherry, enquanto que os clones que apresentam aproximadamente 1100pb tiveram.

Dentre os 16 clones genotipados, três clones foram positivos para o inserto (A Figura 23 mostra um deles – C7), e o sequenciamento destes fragmentos confirmou a inserção do gene repórter no genoma (Figura 24). Entretanto, foi observado que a edição do genoma inseriu deleções adicionais às desejadas, em sequências de nucleotídeos nos três clones. Um trabalho minucioso de deleções de alguns nucleotídeos do COUP-TFII mostrou que existem quatro aminoácidos que são essenciais para a ação repressora de COUP-TFII⁵¹ em alguns contextos, localizados justamente na região que foi deletada em todos os clones (nucleotídeos destacados em vermelho na Figura 24).

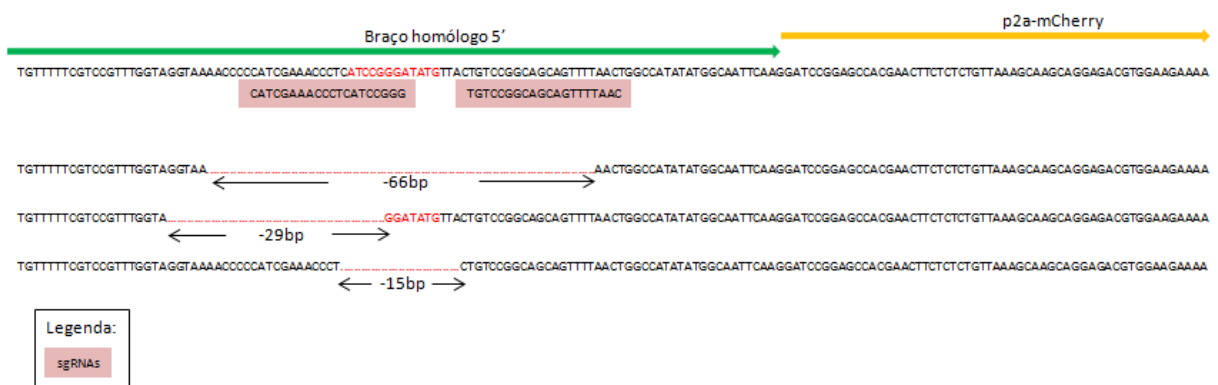


Figura 24: Resultado do sequenciamento de três clones mostrando que houve a inserção do p2a-

mCherry, mas também ocorreu deleção de nucleotídeos nos três clones. Em vermelho estão destacados os códons que dão origem à quatro aminoácidos essenciais para a ação repressora do COUP-TFII.

Uma possível explicação é que, após ter ocorrido a primeira quebra do DNA no sítio esperado e o reparo homólogo com o fragmento doador da sequência, a enzima Cas9n, mantendo a sua capacidade de quebra de DNA e na presença dos sgRNAs a direcionando para este sítio, realizou novas quebras nos alelos já editados. Acreditamos que este tenha sido o ocorrido pois notamos, após análises das sequências antes e depois da edição, que os sítios de ligação dos guias são mantidos na mesma posição “tipo selvagem” na nova sequência sendo inserida nesta região. Mesmo com as deleções, os dois clones que não sofreram frameshift (-66pb e -15pb) foram diferenciados e depois as células passaram no citômetro para detectar fluorescência vermelha e outras 7 fluorescências, mas em nenhum dos filtros foram detectadas alterações na quantidade de fluorescência (Figura 25). Também foi realizada a imunohistoquímica para dois clones diferentes de Coup-TFII-mCherry (Figura 26B e C) para o gene do Coup-TFII (filtro verde – 488) e analisado em filtros que permitem a detecção de mCherry (555 e 647). Não houveram expressões diferenciadas nos filtros para detectar mCherry quando comparadas às células selvagens (Tg2a – Figura 26A) e aos controles negativos (células que não receberam o anticorpo primário – Figura 26D e E).

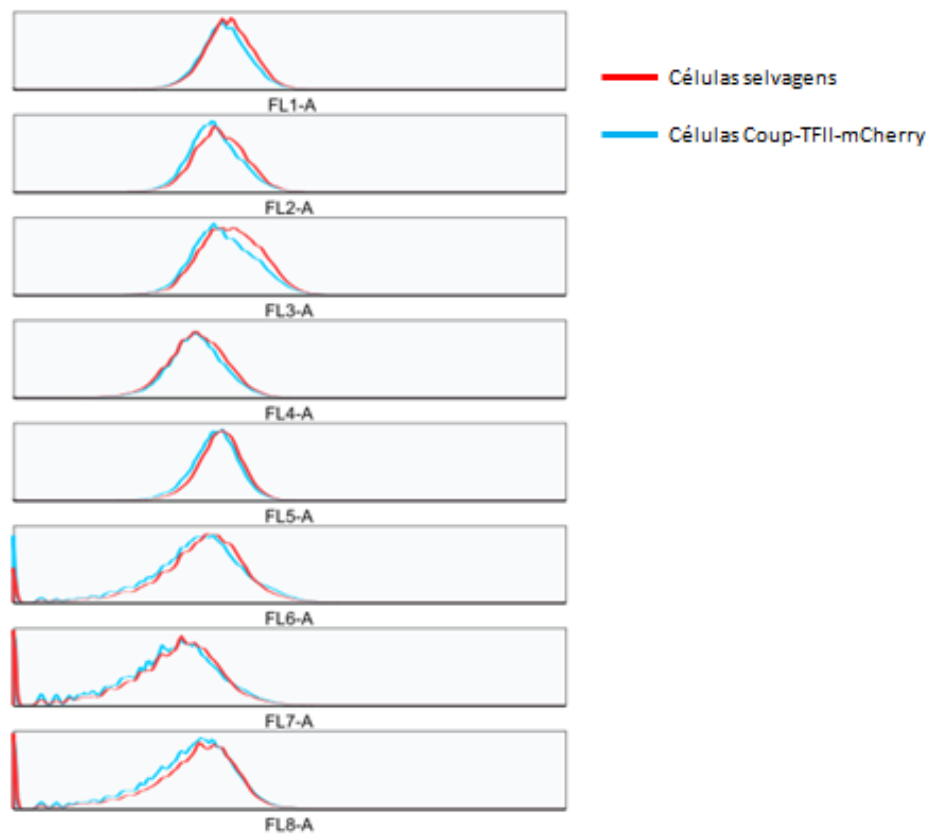


Figura 25: Gráficos gerados no citômetro para 8 filtros diferentes com o intuito de ver se há expressão de mCherry.

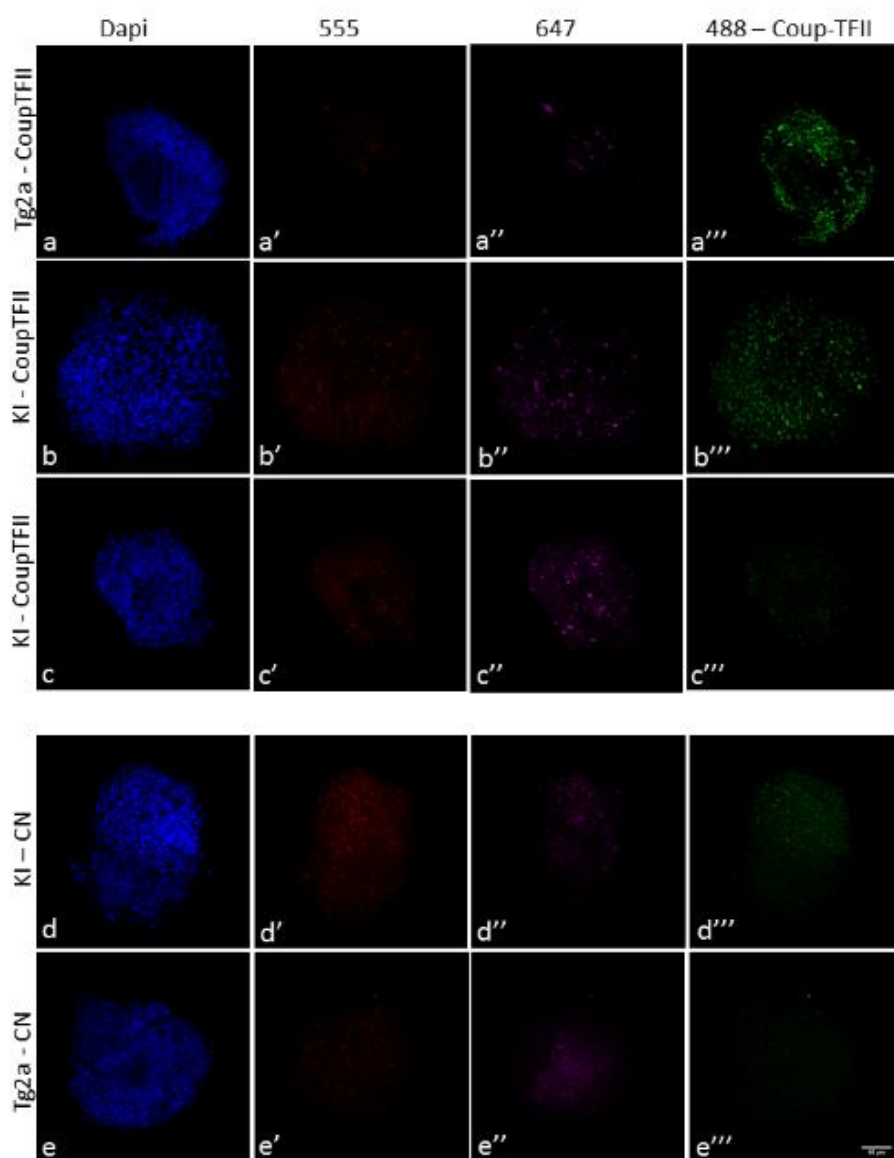


Figura 26: Imunohistoquímica para COUP-TFII em CEs diferenciados por 10 dias a partir de células selvagens e células repórteres mCherry. A- Células selvagens; B – Células Coup-TFII-mCherry; C - Células Coup-TFII-mCherry; D - Células Coup-TFII-mCherry sem o anticorpo primário; E – Células selvagens sem o anticorpo primário.

Para evitar deleções adicionais, foram projetados novos guias para que os sítios de complementariedade tenham distâncias diferentes no alelo tipo selvagem e no alelo editado. Para isso, o guia que está na região do braço homólogo 5' (sgRNA 1 – Tabela 1) foi mantido, sendo desenhando um novo guia na região do braço homólogo 3' (sgRNA 24 – Tabela 1). No alelo tipo selvagem de Coup-TFII, a distância entre os guias é de 80pb, distância que permite que a quebra de fita simples de DNA pela Cas9n em cada sítio resulte em uma quebra de fita dupla nesta região⁶⁷. Após o primeiro ciclo de deleção da Cas9n e inserção do gene repórter, os

sítios de ligação dos guias passam a estar aproximadamente 880pb distantes um do outro. Mesmo que a Cas9n ainda corte uma fita do DNA em cada um dos sítios dos guias, isto não resultará na quebra da dupla fita de DNA⁶⁷, não ocorrendo portanto as deleções de nucleotídeos.

6) Geração de linhagens de CTE nocaute para COUP-TFI por CRISPR/Cas9

Como alguns trabalhos já demonstraram que há uma sobreposição entre COUP-TFI e COUP-TFII onde um pode compensar parte da função do outro⁶⁸, as células foram transfectadas com os sgRNAs 26 e 27 (Tabela 1) para que gerasse o nocaute do COUP-TFI.

Para a geração do nocaute do COUP-TFI, os guias utilizados causaram a quebra da dupla fita de DNA no início do segundo exon, que é um exon comum a todas as variantes de transcrito.

5. CONCLUSÃO

Este projeto permitiu o estabelecimento da técnica de engenharia genômica CRISPR/Cas9 no laboratório, além de contribuir significativamente para o progresso da principal linha de pesquisa do laboratório, a compreensão da função de Coup-TFII nas primeiras tomadas de decisão de células-tronco embrionárias. Com as linhagens celulares obtidas neste projeto, iniciou-se a etapa de análises funcionais do projeto, até então marcada por tentativas frustradas por meio de silenciamento gênico. Por estar encarregada da viabilização de ferramentas funcionais da linha de pesquisa do laboratório, tive uma inserção intelectual e técnica em todos os demais projetos do laboratório. Além disso, os ganhos obtidos durante a experiência no MIT foram direta e prontamente transferidos ao grupo de pesquisa, através de otimizações de técnicas de biologia molecular e cultura de células. O impacto dos avanços obtidos neste projeto serão evidenciados através dos novos rumos possibilitados por eles, mas já nos permitiu aumentar a rede de colaborações a cerca destas tecnologias e sermos reconhecidos como um dos primeiros laboratórios a estabelecer e utilizar a técnica no país.

Vale ressaltar que para a geração da primeira linhagem nocaute foi necessário aproximadamente um ano e meio de trabalho. As demais linhagens hoje já em análise no laboratório foram obtidas com três meses. Este fato demonstra que além de possibilitar a geração de linhagens de células-tronco embrionárias com o genoma editado, este projeto capacitou o laboratório para realizar nocautes gênicos de forma rotineira.

Este projeto também me possibilitou trabalhar por seis meses durante o mestrado no laboratório do Dr. Rudolf Jaenisch, do Instituto Whitehead, pertencente ao Massachusetts Institute of Technology (MIT). Em uma visita em seu laboratório, percebemos que tanto o tema quanto a abordagem experimental do meu projeto de mestrado eram muito semelhantes a um projeto de pesquisa sendo conduzido por uma aluna de doutorado do Dr. Jaenisch. Visando a experiência de trabalhar em um dos laboratórios mais competitivos e produtivos do mundo, podendo contribuir com o que havia aprendido no laboratório e fazer parte do grupo de pesquisa do Dr. Jaenisch eu me dediquei integralmente ao referido projeto de pesquisa pelos seis

meses de estágio. Nesta ocasião, aprimorei meus conhecimentos em CRISPR/Cas9, gerando novas linhagens nocaute para genes de interesse do laboratório, como também entrei em contato com técnicas capazes de quantificar uma única molécula de mRNA dentro de uma única célula (sm FISH-RNA), fazendo o trabalho e o entendimento de rede de regulação gênicas mais preciso e eficaz. Este meu engajamento resultou no convite para ser co-autora de um trabalho que foi recentemente submetido para a revista Cell Stem Cell (manuscrito submetido: anexo 1).

6. REFERÊNCIAS

1. Tachibana, M., Epigenetic regulation of mammalian sex determination. *J Med Invest* **2015**, 62 (1-2), 19-23.
2. Kalkan, T.; Smith, A., Mapping the route from naive pluripotency to lineage specification. *Philos Trans R Soc Lond B Biol Sci* **2014**, 369 (1657).
3. Nichols, J.; Smith, A., Naive and primed pluripotent states. *Cell Stem Cell* **2009**, 4 (6), 487-92.
4. Edlund, T.; Jessell, T. M., Progression from extrinsic to intrinsic signaling in cell fate specification: a view from the nervous system. *Cell* **1999**, 96 (2), 211-24.
5. Jaenisch, R.; Bird, A., Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. *Nat Genet* **2003**, 33 Suppl, 245-54.
6. Bedzhov, I.; Graham, S. J.; Leung, C. Y.; Zernicka-Goetz, M., Developmental plasticity, cell fate specification and morphogenesis in the early mouse embryo. *Philos Trans R Soc Lond B Biol Sci* **2014**, 369 (1657).
7. Lajtha, L. G., Stem cell concepts. *Differentiation* **1979**, 14 (1-2), 23-34.
8. Potten, C. S.; Lajtha, L. G., Stem cells versus stem lines. *Ann N Y Acad Sci* **1982**, 397, 49-61.
9. Andrews, P. W., From teratocarcinomas to embryonic stem cells. *Philos Trans R Soc Lond B Biol Sci* **2002**, 357 (1420), 405-17.
10. Pierce, G. B., Neoplasms, differentiations and mutations. *Am J Pathol* **1974**, 77 (1), 103-118.
11. Solter, D., From teratocarcinomas to embryonic stem cells and beyond: a history of embryonic stem cell research. *Nat Rev Genet* **2006**, 7 (4), 319-27.
12. Stevens, L. C.; Little, C. C., Spontaneous Testicular Teratomas in an Inbred Strain of Mice. *Proc Natl Acad Sci U S A* **1954**, 40 (11), 1080-7.
13. STEVENS, L. C., EXPERIMENTAL PRODUCTION OF TESTICULAR TERATOMAS IN MICE. *Proc Natl Acad Sci U S A* **1964**, 52, 654-61.
14. KLEINSMITH, L. J.; PIERCE, G. B., MULTIPOTENTIALITY OF SINGLE EMBRYONAL CARCINOMA CELLS. *Cancer Res* **1964**, 24, 1544-51.
15. PIERCE, G. B.; VERNEY, E. L., An in vitro and in vivo study of differentiation in teratocarcinomas. *Cancer* **1961**, 14, 1017-29.
16. Evans, M. J., The isolation and properties of a clonal tissue culture strain of pluripotent mouse teratoma cells. *J Embryol Exp Morphol* **1972**, 28 (1), 163-76.
17. Martin, G. R.; Evans, M. J., Differentiation of clonal lines of teratocarcinoma cells: formation of embryoid bodies in vitro. *Proc Natl Acad Sci U S A* **1975**, 72 (4), 1441-5.
18. Martin, G. R., Teratocarcinomas as a model system for the study of embryogenesis and neoplasia. *Cell* **1975**, 5 (3), 229-43.

19. Brinster, R. L., The effect of cells transferred into the mouse blastocyst on subsequent development. *J Exp Med* **1974**, *140* (4), 1049-56.
20. Rossant, J.; McBurney, M. W., The developmental potential of a euploid male teratocarcinoma cell line after blastocyst injection. *J Embryol Exp Morphol* **1982**, *70*, 99-112.
21. Blelloch, R. H.; Hochedlinger, K.; Yamada, Y.; Brennan, C.; Kim, M.; Mintz, B.; Chin, L.; Jaenisch, R., Nuclear cloning of embryonal carcinoma cells. *Proc Natl Acad Sci U S A* **2004**, *101* (39), 13985-90.
22. Li, M.; Belmonte, J. C., Ground rules of the pluripotency gene regulatory network. *Nat Rev Genet* **2017**.
23. Thomson, J. A.; Itskovitz-Eldor, J.; Shapiro, S. S.; Waknitz, M. A.; Swiergiel, J. J.; Marshall, V. S.; Jones, J. M., Embryonic stem cell lines derived from human blastocysts. *Science* **1998**, *282* (5391), 1145-7.
24. Evans, M. J.; Kaufman, M. H., Establishment in culture of pluripotential cells from mouse embryos. *Nature* **1981**, *292* (5819), 154-6.
25. Martin, G. R., Isolation of a pluripotent cell line from early mouse embryos cultured in medium conditioned by teratocarcinoma stem cells. *Proc Natl Acad Sci U S A* **1981**, *78* (12), 7634-8.
26. Wu, J.; Yamauchi, T.; Izpisua Belmonte, J. C., An overview of mammalian pluripotency. *Development* **2016**, *143* (10), 1644-8.
27. Williams, R. L.; Hilton, D. J.; Pease, S.; Willson, T. A.; Stewart, C. L.; Gearing, D. P.; Wagner, E. F.; Metcalf, D.; Nicola, N. A.; Gough, N. M., Myeloid leukaemia inhibitory factor maintains the developmental potential of embryonic stem cells. *Nature* **1988**, *336* (6200), 684-7.
28. Doetschman, T. C.; Eistetter, H.; Katz, M.; Schmidt, W.; Kemler, R., The in vitro development of blastocyst-derived embryonic stem cell lines: formation of visceral yolk sac, blood islands and myocardium. *J Embryol Exp Morphol* **1985**, *87*, 27-45.
29. Nichols, J.; Zevnik, B.; Anastassiadis, K.; Niwa, H.; Klewe-Nebenius, D.; Chambers, I.; Schöler, H.; Smith, A., Formation of pluripotent stem cells in the mammalian embryo depends on the POU transcription factor Oct4. *Cell* **1998**, *95* (3), 379-91.
30. Niwa, H.; Miyazaki, J.; Smith, A. G., Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* **2000**, *24* (4), 372-6.
31. Avilion, A. A.; Nicolis, S. K.; Pevny, L. H.; Perez, L.; Vivian, N.; Lovell-Badge, R., Multipotent cell lineages in early mouse development depend on SOX2 function. *Genes Dev* **2003**, *17* (1), 126-40.
32. Masui, S.; Nakatake, Y.; Toyooka, Y.; Shimosato, D.; Yagi, R.; Takahashi, K.; Okochi, H.; Okuda, A.; Matoba, R.; Sharov, A. A.; Ko, M. S.; Niwa, H., Pluripotency governed by Sox2 via regulation of Oct3/4 expression in mouse embryonic stem cells. *Nat Cell Biol* **2007**, *9* (6), 625-35.

33. Thomson, M.; Liu, S. J.; Zou, L. N.; Smith, Z.; Meissner, A.; Ramanathan, S., Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell* **2011**, *145* (6), 875-89.
34. Chambers, I.; Silva, J.; Colby, D.; Nichols, J.; Nijmeijer, B.; Robertson, M.; Vrana, J.; Jones, K.; Grotewold, L.; Smith, A., Nanog safeguards pluripotency and mediates germline development. *Nature* **2007**, *450* (7173), 1230-4.
35. Mitsui, K.; Tokuzawa, Y.; Itoh, H.; Segawa, K.; Murakami, M.; Takahashi, K.; Maruyama, M.; Maeda, M.; Yamanaka, S., The homeoprotein Nanog is required for maintenance of pluripotency in mouse epiblast and ES cells. *Cell* **2003**, *113* (5), 631-42.
36. Young, R. A., Control of the embryonic stem cell state. *Cell* **2011**, *144* (6), 940-54.
37. Boyer, L. A.; Lee, T. I.; Cole, M. F.; Johnstone, S. E.; Levine, S. S.; Zucker, J. P.; Guenther, M. G.; Kumar, R. M.; Murray, H. L.; Jenner, R. G.; Gifford, D. K.; Melton, D. A.; Jaenisch, R.; Young, R. A., Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* **2005**, *122* (6), 947-56.
38. Hackett, J. A.; Surani, M. A., Regulatory principles of pluripotency: from the ground state up. *Cell Stem Cell* **2014**, *15* (4), 416-30.
39. Tiscornia, G.; Izpisua Belmonte, J. C., MicroRNAs in embryonic stem cell function and fate. *Genes Dev* **2010**, *24* (24), 2732-41.
40. Gangaraju, V. K.; Lin, H., MicroRNAs: key regulators of stem cells. *Nat Rev Mol Cell Biol* **2009**, *10* (2), 116-25.
41. Houbaviy, H. B.; Murray, M. F.; Sharp, P. A., Embryonic stem cell-specific MicroRNAs. *Dev Cell* **2003**, *5* (2), 351-8.
42. Rosa, A.; Brivanlou, A. H., A regulatory circuitry comprised of miR-302 and the transcription factors OCT4 and NR2F2 regulates human embryonic stem cell differentiation. *EMBO J* **2011**, *30* (2), 237-48.
43. Hu, S.; Wilson, K. D.; Ghosh, Z.; Han, L.; Wang, Y.; Lan, F.; Ransohoff, K. J.; BurrIDGE, P.; Wu, J. C., MicroRNA-302 increases reprogramming efficiency via repression of NR2F2. *Stem Cells* **2013**, *31* (2), 259-68.
44. Qiu, Y.; Krishnan, V.; Zeng, Z.; Gilbert, D. J.; Copeland, N. G.; Gibson, L.; Yang-Feng, T.; Jenkins, N. A.; Tsai, M. J.; Tsai, S. Y., Isolation, characterization, and chromosomal localization of mouse and human COUP-TF I and II genes. *Genomics* **1995**, *29* (1), 240-6.
45. Wu, S. P.; Yu, C. T.; Tsai, S. Y.; Tsai, M. J., Choose your destiny: Make a cell fate decision with COUP-TFII. *J Steroid Biochem Mol Biol* **2016**, *157*, 7-12.
46. Le Guével, R.; Oger, F.; Martinez-Jimenez, C. P.; Bizot, M.; Gheeraert, C.; Firmin, F.; Ploton, M.; Kretova, M.; Paliarne, G.; Staels, B.; Barath, P.; Talianidis, I.; Lefebvre, P.; Eeckhoutte, J.; Salbert, G., Inactivation of the Nuclear Orphan Receptor COUP-TFII by Small Chemicals. *ACS Chem Biol* **2017**.

47. Cooney, A. J.; Leng, X.; Tsai, S. Y.; O'Malley, B. W.; Tsai, M. J., Multiple mechanisms of chicken ovalbumin upstream promoter transcription factor-dependent repression of transactivation by the vitamin D, thyroid hormone, and retinoic acid receptors. *J Biol Chem* **1993**, *268* (6), 4152-60.
48. Tsai, S. Y.; Tsai, M. J., Chick ovalbumin upstream promoter-transcription factors (COUP-TFs): coming of age. *Endocr Rev* **1997**, *18* (2), 229-40.
49. Xie, X.; Tang, K.; Yu, C. T.; Tsai, S. Y.; Tsai, M. J., Regulatory potential of COUP-TFs in development: stem/progenitor cells. *Semin Cell Dev Biol* **2013**, *24* (10-12), 687-93.
50. Lin, F. J.; Qin, J.; Tang, K.; Tsai, S. Y.; Tsai, M. J., Coup d'Etat: an orphan takes control. *Endocr Rev* **2011**, *32* (3), 404-21.
51. Achatz, G.; Hölzl, B.; Speckmayer, R.; Hauser, C.; Sandhofer, F.; Paulweber, B., Functional domains of the human orphan receptor ARP-1/COUP-TFII involved in active repression and transrepression. *Mol Cell Biol* **1997**, *17* (9), 4914-32.
52. Qin, J.; Chen, X.; Yu-Lee, L. Y.; Tsai, M. J.; Tsai, S. Y., Nuclear receptor COUP-TFII controls pancreatic islet tumor angiogenesis by regulating vascular endothelial growth factor/vascular endothelial growth factor receptor-2 signaling. *Cancer Res* **2010**, *70* (21), 8812-21.
53. Pereira, F. A.; Qiu, Y.; Zhou, G.; Tsai, M. J.; Tsai, S. Y., The orphan nuclear receptor COUP-TFII is required for angiogenesis and heart development. *Genes Dev* **1999**, *13* (8), 1037-49.
54. Takahashi, K.; Yamanaka, S., Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* **2006**, *126* (4), 663-76.
55. Plath, K.; Lowry, W. E., Progress in understanding reprogramming to the induced pluripotent state. *Nat Rev Genet* **2011**, *12* (4), 253-65.
56. Robinton, D. A.; Daley, G. Q., The promise of induced pluripotent stem cells in research and therapy. *Nature* **2012**, *481* (7381), 295-305.
57. Bredenoord, A. L.; Clevers, H.; Knoblich, J. A., Human tissues in a dish: The research and ethical implications of organoid technology. *Science* **2017**, *355* (6322).
58. You, L. R.; Lin, F. J.; Lee, C. T.; DeMayo, F. J.; Tsai, M. J.; Tsai, S. Y., Suppression of Notch signalling by the COUP-TFII transcription factor regulates vein identity. *Nature* **2005**, *435* (7038), 98-104.
59. Wu, S. P.; Cheng, C. M.; Lanz, R. B.; Wang, T.; Respress, J. L.; Ather, S.; Chen, W.; Tsai, S. J.; Wehrens, X. H.; Tsai, M. J.; Tsai, S. Y., Atrial identity is determined by a COUP-TFII regulatory network. *Dev Cell* **2013**, *25* (4), 417-26.
60. Xie, X.; Qin, J.; Lin, S. H.; Tsai, S. Y.; Tsai, M. J., Nuclear receptor chicken ovalbumin upstream promoter-transcription factor II (COUP-TFII) modulates mesenchymal cell commitment and differentiation. *Proc Natl Acad Sci U S A* **2011**, *108* (36), 14843-8.

61. Naka, H.; Nakamura, S.; Shimazaki, T.; Okano, H., Requirement for COUP-TFI and II in the temporal specification of neural stem cells in CNS development. *Nat Neurosci* **2008**, *11* (9), 1014-23.
62. Ran, F. A.; Hsu, P. D.; Wright, J.; Agarwala, V.; Scott, D. A.; Zhang, F., Genome engineering using the CRISPR-Cas9 system. *Nat Protoc* **2013**, *8* (11), 2281-308.
63. Yamazaki, T.; Suehiro, J.; Miyazaki, H.; Minami, T.; Kodama, T.; Miyazono, K.; Watabe, T., The COUP-TFII variant lacking a DNA-binding domain inhibits the activation of the Cyp7a1 promoter through physical interaction with COUP-TFII. *Biochem J* **2013**, *452* (2), 345-57.
64. Hooper, M.; Hardy, K.; Handyside, A.; Hunter, S.; Monk, M., HPRT-deficient (Lesch-Nyhan) mouse embryos derived from germline colonization by cultured cells. *Nature* **1987**, *326* (6110), 292-5.
65. Siegele, D. A.; Hu, J. C., Gene expression from plasmids containing the araBAD promoter at subsaturating inducer concentrations represents mixed populations. *Proc Natl Acad Sci U S A* **1997**, *94* (15), 8168-72.
66. Kopp, J. L.; Ormsbee, B. D.; Desler, M.; Rizzino, A., Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem Cells* **2008**, *26* (4), 903-11.
67. Ran, F. A.; Hsu, P. D.; Lin, C. Y.; Gootenberg, J. S.; Konermann, S.; Trevino, A. E.; Scott, D. A.; Inoue, A.; Matoba, S.; Zhang, Y.; Zhang, F., Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. *Cell* **2013**, *154* (6), 1380-9.
68. Tang, K.; Xie, X.; Park, J. I.; Jamrich, M.; Tsai, S.; Tsai, M. J., COUP-TFs regulate eye development by controlling factors essential for optic vesicle morphogenesis. *Development* **2010**, *137* (5), 725-34.

7. ANEXOS

ANEXO 1

An interconnected network of pluripotency and lineage regulators controls exit from pluripotency

Johanna Goldmann¹, Denes Hnisz¹, Amanda Araujo Gomes Ferreira², Stefan Semrau³, Richard Young^{1,4} and Rudolf Jaenisch^{1,4}

Affiliations:

¹ Whitehead Institute for Biomedical Research, 9 Cambridge Center, Cambridge, MA 02142, USA

² Department of Biochemistry and Tissue Biology, Institute of Biology, University of Campinas, UNICAMP, R. Charles Darwin, CP 6109, Campinas, SP 13083-863, Brazil.

³ Leiden Institute Of Physics, 2333 CC Leiden, The Netherlands

⁴ Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, 02139, USA

*Correspondence to: jaenisch@wi.mit.edu

Summary

The pluripotent state of embryonic stem cells (ESCs) is controlled by a network of transcriptional regulators. The downregulation of this network is thought to occur prior to and to be a prerequisite of differentiation. In this model pluripotency and lineage regulators are thought to be independent. Here we use single cell analysis to propose an alternative model in which the pluripotency and lineage regulators act as an interconnected network to coordinate the exit from pluripotency. We demonstrate that upregulation of lineage regulators precedes the downregulation of the pluripotency network. Genetic disruption of lineage regulators trapped ESCs in a pluripotent state under conditions promoting differentiation, and pluripotency regulators were

preferentially lost in cells differentiating towards different lineages. These results demonstrate that lineage regulators are necessary for the downregulation of the pluripotency network, and suggest that pluripotency and lineage regulators form an interconnected network to control cell fate transitions.

Introduction

Embryonic stem cells (ESCs) derived from the inner cell mass (ICM) of the developing mammalian embryo are defined by their pluripotency i.e. the ability to differentiate into all somatic cell lineages, their self-renewal capacity and a gene expression profile that is similar to that in cleavage embryos (De Los Angeles et al., 2015; Hackett and Surani, 2014; Martello and Smith, 2014). Due to their ability to differentiate into virtually any cell type, ESCs hold enormous potential for disease modeling and regenerative medicine (Ebert et al., 2012; Saha and Jaenisch, 2009; Svendsen, 2013; Tabar and Studer, 2014). While pluripotent cells have been subject to extensive investigation (Boroviak and Nichols, 2014; Chambers and Tomlinson, 2009; Chen et al., 2008; De Los Angeles et al., 2015; Dunn et al., 2014; Jaenisch and Young, 2008; Loh et al., 2015; Orkin et al., 2008; Wang et al., 2006), the nature of how cells exit pluripotency is less well understood. The importance of understanding this transition is underscored by the recognition that the potency of cells used for regenerative medicine represents an intrinsic risk factor of tumorigenic potential in a clinical setting (Ben-David and Benvenisty, 2011; Kawamata et al., 2015; Peterson et al., 2016; Saha and Jaenisch, 2009). Improved understanding of lineage commitment may help better understanding embryonic development, and improve stem cell differentiation strategies for regenerative medicine.

The maintenance of pluripotency in ESCs and transition to differentiation has been proposed to be controlled by a set of cell-specific master transcription factors that form cell-specific gene regulatory networks (Boyer et al., 2005; Fagnocchi et al., 2016; Festuccia et al., 2013; Filipczyk et al., 2015; Huang et al., 2015; Orkin et al., 2008). In ESCs, these master transcription factors include Oct4, Sox2 and Nanog as well as Esrrb (Mitsui et al., 2003; Radzishchanskaya and Silva, 2014; Rizzino and Wuebben, 2016; Weinberger et al., 2016). The exit from pluripotency and subsequent lineage commitment

is accompanied by both the loss of the pluripotency network and the establishment of an alternate, lineage-specific transcriptional network which includes master transcription factors that act as lineage specifiers such as Pax6 for the ectodermal lineage (Zhang et al., 2010) and Gata6 for the endodermal lineage (Capo-Chichi et al., 2005).

The current model on how ESCs transition out of pluripotency suggests that the gene regulatory networks of ESCs and differentiated cells are self-contained, and that the downregulation of the pluripotency network precedes the upregulation of the lineage specification network. This model has been primarily supported by genome-wide studies in cell populations that did not detect high expression of lineage factors during or prior to the exit of pluripotency (Boroviak and Nichols, 2014; Hayashi et al., 2011; Kalkan and Smith, 2014; Lawson et al., 1991). Several findings however appear inconsistent with this model. For example, overexpression of pluripotency factors such as Oct4 and Sox2 can activate lineage networks, rather than stabilizing the pluripotent state. (Kopp et al., 2008; Lawson et al., 1991; Niwa et al., 2000; Thomson et al., 2011). Furthermore, overexpression of lineage regulators such as Pax6 or Gata6 in differentiated cells together with pluripotency factors leads to the reprogramming to a pluripotent cell state (Montserrat et al., 2013; Shu and Deng, 2013; Shu et al., 2013; Takahashi and Yamanaka, 2006) suggesting an alternative model, whereby pluripotency and lineage regulators have interdependent roles in exiting pluripotency.

Here we test the model that pluripotency and lineage regulators have interdependent roles during the exit from pluripotency by using single cell and single molecule technologies. We found that immediately after the signaling environment maintaining pluripotency is disrupted, pluripotency and lineage regulators are co-expressed, and that expression of the lineage regulators is necessary for the downregulation of the pluripotency network. We further show that the lineage regulators have a direct effect on the downregulation of specific pluripotency regulators and that this specificity is predictive of lineage choice. Our results indicate that pluripotency and lineages regulators have interdependent roles and form an interconnected network to control the exit from pluripotency towards lineage specification.

Results

Single molecule RNA FISH to study the exit from pluripotency at a single cell resolution

To test the model that pluripotency and lineage regulators have interdependent roles during early cell state transitions we set up an ESC differentiation system using retinoic acid (RA) as the differentiation inducer. RA was chosen, as it is a well-studied physiological signal that has been implicated in the differentiation of multiple lineages in the developing embryo (Cho et al., 2012; D'Aniello and Waxman, 2015; Mammadova et al., 2016; Rhinn and Dolle, 2012; Schlupf and Steinbeisser, 2014). Embryonic stem cells (ESC) were grown in 2i culture conditions, which minimize transcriptional heterogeneity (Ying et al., 2008) and were then exposed to RA. Samples were taken at regular (6-24h) intervals for 96 hours and analyzed by single molecule mRNA FISH (sm-RNA FISH) through selected marker genes (Figure 1A). We chose sm-RNA FISH as a single cell readout as it allows the quantitative detection of transcripts expressed at low levels in single cells with robust detection of 1-3 transcripts and up to 3 different genes per individual cell (Buganim et al., 2013; Faddah et al., 2013; Hansen and van Oudenaarden, 2013; Kumar et al., 2014; Raj et al., 2008).

Quantification of single transcripts in ESCs revealed robust expression of genes implicated in pluripotency, and characteristic changes in the expression of known lineage regulators in differentiating ESCs. We found that ESCs cultured in 2i express high transcript levels of the core pluripotency factors Oct4 (389.4 ± 114.3), Sox2 (188.5 ± 64.0) and Nanog (130.6 ± 57.4) (Figure 1B, Supplementary Figure 1A,C) (Boroviak and Nichols, 2014) as well as high levels of Esrrb (113.3 ± 46.8) (Supplementary Figure 1A,C). As a control, we detected no significant expression of the lineage regulators Pax6 (0.023 ± 0.28) and Gata6 (0.009 ± 0.095) in ESCs grown in 2i (Supplementary Figure 1B,D). After differentiating cells with RA for 96 hours, the expression of pluripotency regulators was significantly reduced (Supplementary Figure 1E) and expression of the lineage regulators Pax6 (7.08 ± 12.15) and Gata6 (68.55 ± 91.77) was detected (Figure 1C). The average number of transcripts detected by sm-RNA FISH in the whole population of cells correlated with the expression levels measured by RNA-Seq (Supplementary Figure 1F,G). Quantification of the average number of

transcripts of selected pluripotency regulators during RA in over 300 single cells indicated that *Esrrb* was downregulated more rapidly than *Oct4* and *Sox2* following exposure to RA ([Supplementary Figure 1H,I](#)) consistent with previously published studies investigating populations of cells (Leeb et al., 2014). These results indicate that the single molecule RNA FISH-based detection system captures essential previously described features of ESCs and differentiating ESCs.

Resolution of cellular subpopulations

Single molecule RNA FISH is capable of resolving individual cells within cell populations, and thus enables the identification of subpopulations of cells during ESC differentiation that cannot be observed when investigating populations of cells. As expected analysis of *Oct4* and *Sox2* transcripts revealed a gradual downregulation of both genes during differentiation ([Figure 1D](#)). Furthermore RNA-FISH revealed that expression of *Sox2* and *Oct4* were highly correlated and no subpopulations preferentially expressing either *Sox2* or *Oct4* could be found ($p = .69$) when cultured in 2i ([Figure 1D,E](#)). In contrast after 96h of RA exposure ESCs showed a significant increase of the average transcript numbers of *Gata6*, an essential regulator of the endoderm lineage, and *Pax6*, and essential regulator of the ectoderm lineage consistent with previously published studies investigating populations of cells exposed to RA (Capo-Chichi et al., 2005), and with our own RNA-Seq data of populations of cells ([Figure 1F,G](#) and [Supplementary Figure 1G](#)). Importantly, resolution of single cells revealed that *Gata6* and *Pax6* expression was mutually exclusive, i.e. cells that expressed *Pax6* did not expressed *Gata6*, and cells that expressed *Gata6* did not express *Pax6* ([Figure 1G](#)), suggesting that two distinct cell subpopulations had formed by 96 hours ($p = -.12$). These results demonstrate that sm-RNA FISH provides a robust and quantitative readout of cell state marker transcripts during ESC differentiation and allows for the characterization of specific subpopulations formed during the differentiation process.

Lineage and pluripotency regulators are co-expressed prior to and during exiting from pluripotency

If pluripotency and lineage regulators act co-dependently to control the transition from pluripotency to lineage commitment, they would need to be co-expressed prior to and during the exit from pluripotency. We tested this notion by defining the point of exiting pluripotency through a functional assay, and quantifying the expression of pluripotency and lineage regulators during defined timepoints. This analysis revealed that both prior to and during the exit from pluripotency a significant percentage of cells co-expressed both pluripotency and lineage regulators.

The exit from pluripotency was defined using a cell survival assay based on the observation that cells that have exited pluripotency are not viable in 2i (Silva et al., 2008). The cells were exposed to RA to induce differentiation into endodermal and ectodermal cells ([Supplementary Figure 2A,B](#)) for various durations and then replated in 2i. The viability of cells was assessed 48 hours after being plated in 2i. In this assay almost 100% of the cells retained pluripotency when exposed to RA for 12 hours or less, while >50% of cells lost viability after 24 hours and >90% cells had lost viability after 48 hours. ([Supplementary Figure 2C](#)). Based on these data we define 24-48h of RA exposure as the window for exiting the pluripotent state, which is consistent with the downregulation of the pluripotency regulators Nanog and Esrrb ([Supplementary Figure 2D](#)).

Sm-RNA FISH revealed co-expression of pluripotency and lineage regulators in a significant subpopulation of single cells prior to the exit from pluripotency. The lineage regulators Pax6 and Gata6 were virtually absent in 2i conditions with only ~2% (13/547) of cells expressing a single transcript of Pax6 and less than 1% (5/547) expressing Gata6 ([Supplementary Figure 1B,D](#) and [Supplementary Figure 2B](#)). After exposure to RA for 6 and 12 hours, we observed rapid upregulation of both Pax6 and Gata6 on the single cell level with sm-RNA Fish compared to cells grown in 2i ([Figure 2A](#)). More than 72% of all cells co-expressed pluripotency and lineage regulators after 6 hours and this percentage rose to 85% of all cells after 12 hours ([Figure 2C](#)). The average expression level of both Gata6 and Pax6 was around 3-6 transcripts per cell after 6-12 hours and the difference in expression compared to the expression of ESCs was highly significant ($p < 0.0001$ for all timepoints compared to 2i) ([Figure 2C](#)). We also observed a population of cells co-expressing all three markers (Pax6, Gata6, Nanog). This observation is in

accordance with recent findings in regard to transient mixed-lineage stages that might be a prerequisite for differentiation by enabling cell fate specification through dynamic instability (Olsson et al., 2016)

To determine if co-expression of lineage and pluripotency regulators continued during the exit from pluripotency we analyzed their co-expression during the exit of pluripotency (24-48h) (Figure 2D-H, Supplementary Figure 2C). This confirmed that co-expression of lineage and pluripotency regulators occurred at all time points during the cell state transition. The percentage of cells co-expressing lineage and pluripotency regulators peaked at 24 hours with 88% and decreasing during the exit to 34% and 13% at 36h and 48h respectively (Figure 2E, Supplementary Figure 2E).

Co-expression of lineage and pluripotency regulators occurs in multiple differentiation systems

To confirm co-expression of pluripotency and lineage regulators within the same cell we used an alternative differentiation system to induce ectodermal but not endodermal cells within 96 hours (Supplementary Figure 3A). After 96h cells expressed high transcript levels of the ectoderm regulator Pax6 (14 +/- 19), while the endoderm regulator Gata6 was hardly expressed (1 +/- 1) (Supplementary Figure 3B,C). The colony formation assay revealed that in this differentiation system cells exited pluripotency between 36-60h with ~50% cells surviving replating after 36h and less than 10% after 60 hours (Supplementary Figure 3D). Based on these data we investigated the co-expression of pluripotency and lineage regulators prior to the exit at 12 and 24 hours. At 12 hours Pax6 was detected with an average expression of 3 transcripts in 70% of the cells while at 24 hours 76 % of the cells expressed Pax6 with an average expression of 11 transcripts per cell confirming upregulation of lineage regulators (Supplementary Figure 3E). The difference in expression compared to the expression of ESCs was highly significant ($p = < 0.0001$ for all timepoints) (Supplementary Figure 3E-G). In contrast Gata6, which was expressed at an average of less than 2 transcripts even at 96 hours, was expressed in fewer than 14% of cells at 12 hours at a level of 0.19 at 12 hours and a level of 2.1 transcripts in a similar percentage of cells (Supplementary Figure 3E-G). These results demonstrate that lineage regulators are co-expressed with pluripotency regulators in single cells prior and during the cell state transition out of pluripotency in two different

ESC differentiation systems.

Lineage regulators are required for the downregulation of pluripotency regulators

If lineage and pluripotency regulators form an interconnected network to control the exit from pluripotency, lineage regulators should have a functional impact on the cell state transition. To test this idea, we generated homozygous mutants for the ectoderm regulator Pax6 and the endoderm regulator Gata6 using CRISPR/Cas9-based genome editing ([Supplementary Figure 3A,B](#)). Wild type and knockout cells were differentiated as previously described using RA and the pluripotent state was assessed through sm-RNA FISH analysis and colony formation assays ([Supplementary Figure 2A,B](#)).

These experiments revealed that in the absence of the lineage regulators Pax6 and Gata6 cells failed to fully downregulate pluripotency regulators even after 96 hours of differentiation with RA. When cultured in 2i both wild type and KO cells expressed the pluripotency regulators Oct4 and Sox2 at comparable levels ([Supplementary Figure 4C](#)). As expected Pax6(-/-) cells showed did not express Pax6 transcripts after 96h of differentiation while Gata6(-/-) did not express Gata6 transcripts ([Supplementary Figure 4D](#)). However smRNA FISH analysis revealed that WT cells had lost expression of Oct4, Sox2, Nanog and Esrrb after 96 hours, indicating that the WT cells downregulated the pluripotency network. In contrast both knockout cell lines maintained expression of Oct4, Sox2 as well as Nanog and Esrrb ([Figure 3A](#)). While WT cells expressed Oct4 at an average of 1 transcripts per cell (+/- 3), Pax6 (-/-) cells expressed Oct4 at an average of 29 transcripts (+/- 60) and Gata6 (-/-) cells at an average of 140.9 (+/- 133) transcripts per cell. Sox2, Nanog and Esrrb showed a similar pattern ([Figure 3B,C](#)). Furthermore less than 3% of WT cells expressed 10 or more transcripts of Oct4 while 38% of Pax6(-/-) and 75% of Gata6(-/-) cells did. Again a similar pattern could be observed for Sox2 and Nanog and Esrrb ([Figure 3C](#)).

To further assess the cell state of both WT and KO cells after 96 hours of differentiation we performed a colony formation assay. Briefly, cells were exposed to RA for 96 hours then replated in standard ESC medium containing serum and LIF. After 5 days cells were harvested and assayed for the expression of Oct4, Sox2 and Nanog. As expected the majority of WT cells had lost expression of all three markers. In contrast both Pax6 and Gata6 retained high levels of expression of all three markers ([Figure](#)

3D,E). Taken together these data indicate that in the absence of the lineage regulators Pax6 and Gata6 cells were neither able to fully downregulate their pluripotency markers nor able to fully exit from pluripotency.

Lineage regulators are required for exit of pluripotency in multiple differentiation systems

Investigating cells in an alternative differentiation system confirmed the requirement for lineage regulators for the downregulation of the pluripotency network. We differentiated Pax6(-/-) towards the ectodermal lineage as previously described (Supplementary Figure 3A-D). This confirmed that Pax6(-/-) cells retained expression of pluripotency regulators even after 96 hours of differentiation towards the ectodermal lineage. After 96h of differentiation Pax6(-/-) maintained high expression of Oct4, Sox2 as well as Nanog and Esrrb (Figure 4B). WT cells expressed Oct4 expression at a level of 10 or more transcripts in less than 10% of wild type cells at an average number of 4.4 (+/- 1) while more than 90% of Pax6(-/-) cells expressed Oct4 at average number of transcripts of 125 (+/- 106) (Figure 4B,C) Differentiating Pax6(-/-) cells also maintained expression of additional pluripotency markers such as Sox2, Esrrb and Nanog at high levels compared to WT cells ($p < 0.001$) (Figure 4B,C)

To further assess the pluripotency cell state we performed colony formation assays on WT and Pax6(-/-) cells passaged under differentiating culture conditions after transferring them to 2i conditions to functionally assess the pluripotent state of these cells. Wild type cells lost their ability to form colonies in 2i after 36 hours of being cultured in the differentiation medium, indicating that beyond that point they have exited pluripotency (Figure 4D). In contrast, Pax6(-/-) cells passaged under ectoderm differentiating conditions maintained their ability to form colonies after transferring them to 2i conditions (Figure 4D), indicating that they maintained their pluripotent state.

Notably the dependency on lineage regulators for exiting pluripotency increased if the missing lineage regulator was required for subsequent lineage commitment. Cells missing the ectodermal lineage regulator Pax6 retained significantly higher levels of pluripotency markers when cultured in conditions favoring ectoderm (Figure 4E). In contrast KO lines for the endodermal regulator Gata6 retained significantly higher levels of pluripotency regulators when differentiated with conditions permitting expression of

endoderm (Figure 4F). These results indicate that the requirement for lineage regulators to exit pluripotency might be linked with subsequent lineage specification. These results demonstrate that the presence of lineage regulators is required for ESCs to fully downregulate the pluripotency network and exit pluripotency. Furthermore the results indicate the exit from pluripotency and lineage specification are linked.

The mode of exiting pluripotency is lineage specific

A link between exiting from pluripotency and lineage specification would argue for one of two models. Lineage regulators could contribute to the exit from pluripotency through unspecific downregulation of all pluripotency regulators. Alternatively different lineage regulators might specifically affect different pluripotency regulators and would play a role in the temporal order in which the pluripotency network is disassembled. The latter model infers that the exit of pluripotency itself is lineage specific and that the order of downregulation of pluripotency regulators itself may be predictive of lineage choice.

To investigate the two models, we made use of the RA based differentiation system, which induces differentiation towards both ectoderm and endoderm in a mutually exclusive manner. We first analyzed the co-expression of Oct4 and Sox2 at single cell resolution throughout the differentiation time course. All cells cultured under 2i conditions co-expressed Sox2 and Oct4 at a stable ratio of Oct4 and Sox2 ($p = .69$) (Figure 5A,B). As the cells exited pluripotency (24h) two subpopulations started to emerge which either expressed higher levels of Oct4 then Sox2 (27%) or higher levels of Sox2 then Oct4 (9%) (Figure 5C,D). At 72h of RA exposure, less than 12% of cells expressed Oct4 and Sox2 at a stable ratio (Figure 5E,F) with expression of Oct4 and Sox2 being anti correlated ($p = -0.039$).

To determine if selective downregulation of pluripotency regulators is directly relevant to subsequent lineage choice, we quantified the expression of Oct4 and Sox2 with Gata6 during differentiation (Figure 5G). We observed that cells expressing high levels of the endodermal marker Gata6 (Extended Experimental Procedures) at 72 hours of differentiation expressed significantly lower levels of Sox2 (average expression 25 +/- 45) then cells expressing low levels or no Gata6 at all (average expression Sox2 of 60 +/- 66) (Figure 5H).

While no Oct4-high or Sox2 high population were present in 2i, within 24 hours the two subpopulations started to emerge. At this point the subpopulations did not yet show a statistically significant difference in their respective levels of either Gata6 or Pax6 (Figure 5I). However, this might be due to the small percentage of cells forming the subpopulations at this timepoint (<23% of total cells). In contrast, after the exit from pluripotency at 72 hours, the Oct4-high subpopulation contained an average of ~70 transcripts of Gata6 (+/- 130) while the Sox2 high subpopulation contained an average of ~10 transcripts (+/- 3) (Figure 5I). The opposite was found for Pax6 (Supplementary Figure 5A-C).

The lineage specific mode of exiting pluripotency is dependent on lineage regulators

If selective downregulation of pluripotency regulators was dependent on the expression of lineage regulators, inactivation of a lineage regulator should be sufficient to perturb this process. Analysis of both mutant cell lines confirmed this prediction. We analyzed the dynamics of Oct4 and Sox2 co-expression in Gata6(-/-) and Pax6(-/-) mutant cells. In the absence of the lineage regulators Gata6 or Pax6 the majority of cells continued to co-express Oct4 and Sox2 even after 72 hours of differentiation (Figure 6A-D). After 72 hours more than 88% of WT cells had preferentially lost either Oct4 or Sox2. In contrast less than 59% of Pax6(-/-) cells and less than 45% of Gata6(-/-) cells showed preferential loss of Oct4 or Sox2 (Figure 6E). This effect was even more pronounced at the onset of exiting pluripotency at 24 hours (Figure 6E). The shift in subpopulations for both Pax6 and Gata6 mutant lines was statistically significant ($p < 0.0001$). These results show that in the absence of lineage regulators cells, continued to co-express Oct4 and Sox2, leaving the cells trapped in a pluripotency-like state.

Discussion

While the transcriptional foundation of pluripotency has been the topic of intense interest over the past two decades, the exit from pluripotency has only recently started to be investigated (Boroviak and Nichols, 2014; Chambers and Tomlinson, 2009; Chen et al., 2008; De Los Angeles et al., 2015; Dunn et al., 2014; Jaenisch and Young, 2008; Loh et

al., 2015; Orkin et al., 2008; Wang et al., 2006). An in depth understanding of the exit from pluripotency is crucial not only for our understanding of developmental biology but also in order to harness the power of ESCs for regenerative medicine. Here we provide an updated model for the transcriptional events underlying this transition. Using highly quantitative single cell analysis we show that prior to and during the exit of pluripotency, critical transcription factors for both pluripotency and lineage specification are co-expressed. We also demonstrate that perturbation of lineage regulators is sufficient to disrupt this transition. Lastly we show that the cells exit pluripotency in a lineage specific manner and that the loss of particular pluripotency regulators is predictive of lineage choice. We suggest a model in which members of both the lineage and the pluripotency transcriptional network act in concert to control cell fate transitions in embryonic cells (Figure 7).

The current model of how ESCs exit pluripotency postulates that the downregulation of pluripotency regulators precedes the upregulation of the lineage specification network (Kalkan and Smith, 2014). This model has primarily been supported through genome-wide studies in cell populations, which do not detect high expression of lineage regulators prior to the exit from pluripotency (Kopp et al., 2008; Lawson et al., 1991; Niwa et al., 2000; Thomson et al., 2011).

Using a highly sensitive single cell technique our results challenge this view. We demonstrate that the lineage regulators Pax6 and Gata6 are expressed in the majority of cells within 6 hours of differentiation and that lineage and pluripotency regulators are co-expressed as the cells exit pluripotency. Supporting these findings is recent work by the Zernicke-Goetz lab demonstrating co-expression of lineage and pluripotency regulators in the early cleavage embryo (Goolam et al., 2016). The co-expression of both pluripotency and lineage networks within a single cell, in vivo as well as in-vitro, provides the first indication that the two cell states and their respective transcriptional networks might be interconnected.

We used CRISPR/Cas9 to generate homozygous knockouts of the lineage regulators Pax6 and Gata6 and demonstrated that loss of either lineage regulator is sufficient to disrupt the exit from pluripotency. Multiple studies have investigated the exit from pluripotency, using disruption of members of the pluripotency network such as Tfe3, Tcf, Pum and Tbx3 (Betschinger et al., 2013; Leeb et al., 2014; Yang et al., 2012).

Hence this transition has only been studied as a direct consequence of the destabilization of the pluripotency network. Our results are consistent with previous data demonstrating that loss of lineage regulators, including Pax6, is required for the exit from pluripotency (Betschinger et al., 2013; Leeb et al., 2014; Yang et al., 2012). Thus, exit from pluripotency is coordinated not only through the destabilization of the pluripotency network but also depends on the establishment of the lineage specific transcriptional networks.

We show that Oct4 and Sox2 are not lost homogenously within the entire cell population as cells transition towards lineage commitment. Instead the subpopulation of cells that first lost Oct4 while maintaining Sox2 expression upregulates Pax6. In contrast, cells that first lost Sox2 while maintaining Oct4 upregulated expression of Gata6. In the absence of the lineage regulators Pax6 or Gata6 cells no longer showed the preferential loss of either Oct4 or Sox2 leaving the cells trapped in a pluripotent-like state. The association of Oct4 with the endodermal fate is also supported by previous studies showing that overexpression of Oct4 is sufficient to destabilize the pluripotency network and do specifically induce endodermal gene expression including Gata6 (Niwa et al., 2000; Shimozaki et al., 2003). Similarly overexpression of Sox2 also leads to the exit from pluripotency but causes differentiation towards ectoderm (Kopp et al., 2008) suggesting that pluripotency factors can act as lineage specifiers. Consistent with this model, Oct4 partners with the endodermal regulator Sox17 while Sox2 partners with Brn2, a well-known ectodermal regulator during differentiation (Aksoy et al., 2013; Lodato et al., 2013).

Our results provide a new conceptual framework to understand the dynamics behind the exit from pluripotency. Acquisition of a committed cell state is not only dependent on the up-regulation of specific lineage regulators but also on expression level of individual pluripotency factors ([Figure 7](#)). By gaining further understanding how the lineage and pluripotency network interact and how the different roles of both networks are controlled we will gain a deeper understanding of the basic principles of early development

Author contributions

J.G. conceived and designed the study, performed the experiments, analyzed the data and

wrote the manuscript. D.H. provided critical input for the design of the study and wrote the manuscript.

A.A. performed part of the smFISH experiments and Tissue Culture. S.S. provided analysis tools for smFISH analysis and established the RA differentiation system. R.J. and R.Y. supervised the design of the study, experiments and manuscript writing.

Acknowledgements

We are grateful to Tony Lee, Jessica Reddy, Michael Seyffert, Thomas Graf, Tian Tian, Camille Sindhu and Thorold Theunissen for critical discussions and reading of the manuscript. J.G. was supported by the Boehringer Ingelheim Foundation as well as a Jerome and Florence Brill Graduate Student Fellowship. D.H. was supported by an Erwin Schrodinger Fellowship (J3490) from the Austrian Science Fund (FWF) and a Margaret and Herman Sokol Postdoctoral Award. S.S. was supported by the Netherlands Organization for Scientific Research (NWO/OCW), as part of the Frontiers of Nanoscience program. A.A. was supported by FAPESP grants (2012/09602-0 and 2015/06732-9). R.J. is an advisor to Stemgent and co-founder of Fate Therapeutics. R.A.Y. was supported by NS088538 and is a founder of Syros Pharmaceuticals and Marauder Therapeutics.

Experimental Procedures

Cell culture

All cell lines (E14) were grown in modified 2i (Extended Experimental Procedures) and passaged every 2-3 days with accutase and plated on gelatin coated tissue culture plates.

Differentiation

Cells were grown in 2i medium for at least 3 passages and plated at a density of 2.5×10^6 cells per 10cm dish and grown overnight. Cells were washed 2x with PBS and the differentiation medium was added. Medium was refreshed after 48 hours. Samples for smFISH

FISH imaging and analysis

We performed FISH imaging and analysis as previously described (Raj et al., 2008).

Briefly, probes were coupled to TMR, Alexa 594 or Cy5 and images taken with a Nikon Ti-E inverted fluorescence microscope and imaged with a 100x oil immersion objective. All images were taken with a Photometrics Pixis 1024 CCD camera and using Metamorph (Molecular Devices). Analysis of FISH data and quantification of dots was performed using custom written Matlab software written by SS.

CRISPR/Cas9 mediated genome editing

Genome editing was performed using CRISPR/Cas9 essentially as previously described (Yang et al., 2013). Briefly, target-specific oligonucleotides were cloned into a plasmid carrying Cas9 and a GFP expression cassette. sgRNAs were designed by manually and screened using the Cell Assay (IDT DNA). Clones were FACS sorted for GFP expression. Cells were genotyped after sorting and sequenced.

Extended Experimental Procedures

Cell culture

Modified 2i contained DMEM/F12 (Life technologies) supplemented with 0.5x N2 supplement, 0.5x B27 supplement, 4mM L-glutamine (Gibco), 20 µg/ml human insulin (Sigma-Aldrich), 1x 100U/ml penicillin/streptomycin (Gibco), 1x MEM Non-Essential Amino Acids (Gibco), 0.1mM 2- Mercaptoethanol (Sigma-Aldrich), 1 µM MEK inhibitor (PD0325901, Stemgent), 3 µM GSK3 inhibitor (CHIR99021, Stemgent) and mouse LIF. Cells were regularly genotyped (Sanger sequencing of the Pax6 and Gata6 locus) and tested for mycoplasma.

Differentiation

For RA based differentiation cells were differentiated with basal N2B27 medium (2i medium without LIF and inhibitors) and 0.25µM of RA (Sigma Alderich). Cells were washed 2x with PBS to remove any inhibitors and LIF and freshly prepared medium containing RA was added. After 48 hours the medium was refreshed with freshly prepare medium containing RA.

For differentiation towards ectoderm were differentiated with basal N2B27 in the absence of RA . Cells were washed 2x with PBS and the medium was refreshed after 48 hours.

Colony formation assay in 2i

Cells were differentiated with basal N2B27 as described above. At various timepoints the cells were replated at a density of 5×10^4 cells/well in a gelatinized 6-well plate in 2i medium. After replating cells in 2i the colonies were grown for 2 additional days and imaged. Remaining colonies were counted manually.

Colony formation assay in Serum and LIF

Cells were differentiated with 0.25 as described above. At various timepoints the cells were replated at a density of 5×10^4 cells/well in a gelatinized 6-well plate in standard ESC medium containing serum and LIF (Knockout DMEM (Thermofisher) supplemented with 10% ES cell screened FBS (Sigma), 1x 100U/ml penicillin/streptomycin (Gibco), 0.1mM 2-Mercaptoethanol (Sigma-Aldrich) and 1000 U/ml mouse LIF). After replating cells in 2i the colonies were grown for 5 additional days. After 5 days cells were fixed and permeabilized for FISH imaging.

smFISH

Prior to smFISH the cells were fixed and permeabilized. Cells were fixed using 4% PFA in medium and incubated for 10 minutes on a shaker. Subsequently cells were permeabilized in 70% ethanol for a minimum of 24 hours. To hybridize cells the cells were plated in chambered cover slides (Nunc Lab Tek) coated with poly-l-lysine.

All probes were designed and labeled as previously described (Raj et al., 2008). All probes were 20 nucleotide probes. The hybridization buffer used for smFISH contained 2 × SSC buffer, 25% or 40% formamide, 10% Dextran Sulphate (Sigma), E. coli tRNA (Sigma), Bovine Serum Albumin (Ambion) and Ribonucleoside Vanadyl Complex (New England Biolabs). 75 ng of the probes were used per 100 µl of hybridization buffer. Probes were hybridized for 16 -18 h at 30°C. washed twice for 30 min at 30°C in a 25% Formamide containing washbuffer with Hoechst 33342.

smFISH microscopy

Cells were imaged in an imaging buffer (1 x PBS, 0.4% Glucose, 100 µg/ml Catalase, 37 µg/ml Glucose Oxidase, and 2 mM Trolox). Imaging was done as previously described

(Semrau et al., 2014) and home-made MATLAB scripts were used for image analysis.

Ratio Analysis

To calculate the ratio of Oct4 and Sox2, the mean expression level of Oct4 and Sox2 was used (Oct4/Sox2) for the exact experiment and timepoint analyzed. For each individual cell then the ratio (Oct4/Sox2) was calculated and classified as 2X ratio (high Sox2 and low Oct4, pink) of the mean ratio or ½ X ratio (low Sox2 and high Oct4, blue).

To calculate Gata6 high and Pax6 high cells the mean expression level of Gata6 or Pax6 was used for the exact experiment and timepoint analyzed. Cells expressing at least the mean expression level + the SEM were considered to be Gata6 or Pax6 high.

Figure Legends:

Figure 1 – Single molecule mRNA Fish based detection system to study the exit from pluripotency at a single cell resolution

- (A) Schematic of the differentiation assay and marker genes used.
- (B) Representative microscopy images (as Z-projection) for the pluripotency regulators Oct4 and Sox2 and scatter plots (with mean + SD) to show transcript distribution in ESCs under 2i conditions. On the microscopy images each diffraction-limited dot corresponds to a single mRNA molecule. All transcript numbers are presented as a mean ± standard deviation (SD). Hoechst staining of nuclei is shown in blue.
- (C) Representative microscopy images for the lineage regulators Pax6 (ectoderm) and Gata6 (endoderm) and scatter plots to show transcript distribution in ESCs after 96 hours of differentiation with RA.
- (D) Dynamics of downregulation of transcript numbers of Sox2 and Oct4 on the single cell level with sm-FISH. Average transcript numbers were normalized using 0,1 normalization (see Extended Materials Methods).
- (E) Representative microscopy image and quantification showing co-expression of Oct4 and Sox2 in 2i. The Spearman Correlation (ρ) coefficient was calculated to be .69
- (F) Dynamics of upregulation of transcript numbers of Pax6 and Gata6 on the single cell level with sm-FISH.
- (G) Representative microscopy image showing anti-correlation of Pax6 and Gata6 after 96 hours and quantification thereof. The Spearman Correlation (ρ) coefficient was calculated to be -.12

Figure S1 – Related to Figure 1

- (A) Representative microscopy images for the pluripotency regulators Nanog and Esrrb under 2i conditions (as Fig.1A)
- (B) Representative microscopy images for the lineage regulators Pax6 and Gata6 under 2i conditions. (as Fig.1A)

- (C) Scatter plots to show transcript distribution of Nanog and Esrrb in mESCs in 2i conditions
- (D) Scatter plots to show transcript distribution of Pax6 and Gata6 in mESCs in 2i conditions
- (E) Barplots showing average number of transcripts of the pluripotency markers Nanog, Esrrb, Sox2 and Oct4 in 2i and after 96 hours of differentiation.
- (F) Comparison of dynamics of Oct4 and Sox2 transcript level assayed by smFISH and RNA-seq
- (G) Comparison of dynamics of Gata6 and Pax6 transcript level assayed by smFISH and RNA-seq
- (H) Dynamics of downregulation of the core pluripotency factors Oct4 and Sox2 as well as the naïve pluripotency factors Klf2 and Esrrb
- (I) Bar plots showing relative expression level compared to 2i after 48 hours of RA exposure

Figure 2 – Lineage regulators and pluripotency regulators are co-expressed prior to and during the exit from pluripotency

- (A) Representative sm-FISH images of Pax6, Gata6 and Nanog expression in 2i and after 6h and 12h of RA exposure.
- (B) Bar graph of percentage of cells expressing either only Nanog or co-expression Nanog and Pax6 or Nanog and Gata6 in 2i and after 6 and 12 hours of RA exposure. Any cells expressing >1 transcript was considered positive.
- (C) Box and Whisker plots to show distribution of transcripts expression levels for Pax6, Gata6 and Nanog in 2i and 6h and 12h of RA exposure. The whisker as shown always represented 2.5th percentile and the 97.5th percentile. The median expression is indicated within the boxplot. Statistical significance was determined through the Welch's t-test. (***) = $p < 0.0001$, (**) = $p < 0.001$, (*) = $p < 0.01$)
- (D) Representative sm-FISH images of Pax6, Gata6 and Nanog expression after 24, 36, and 48 hours of RA exposure.
- (E) Bar graph of percentage of cells expressing either Pax6 or Gata6 transcripts at 24 hours, 36 hours and 48 hours.
- (F) Box and Whisker plots to show distribution of transcripts expression levels for Pax6, Gata6 and Nanog at 24h, 36h and 48h of RA exposure.

Figure S2 - Related to Figure 2

- (A) Representative sm-Fish images for Gata6, Pax6 and Nanog grown in 2i and after 96 hours of RA exposure.
- (B) Bar graphs for the mean transcript levels of Pax6, Gata6 and Nanog in 2i and after 96 hours of differentiation.

- (C) Colony formation assay to determine the temporal window for the exit from pluripotency. Cells were exposed to RA and after differentiation for the indicated amounts of time cells were replated at a defined, clonal density in 2i. Cells were cultured in 2i medium for 2 additional days and colonies were counted. The percentage of cells surviving was calculated compared to cells never exposed to RA.
- (D) Graph of colony formation assay (Figure S2C) overlaid with expression level of *Esrrb* and *Nanog* as assayed by smFISH. Expression levels of *Nanog* and *Esrrb* were normalized relative to 2i.
- (E) Dynamics of cells co-expressing *Nanog* and *Pax6* or *Nanog* and *Gata6* as well as cells positive for only *Nanog* during the differentiation timecourse.

Figure S3 – Related to Figure 2

- (A) Scheme to represent differentiation assay used for the generation of ectodermal cells. Cells were plated in 2i and then differentiated in basal N2B27 medium.
- (B) Bar graphs for the mean transcript levels of *Pax6*, *Gata6* and *Nanog* in 2i and after 96 hours of differentiation with N2B27.
- (C) Representative sm-FISH images of *Pax6*, *Gata6* and *Nanog* expression in 2i and after 96 hours of N2B27 exposure
- (D) Colony formation assay to determine when the exit from pluripotency occurs under conditions inducing ectoderm.
- (E) Representative sm-FISH images of *Pax6*, *Gata6* and *Nanog* expression after 12h, 24h and 48 hours of differentiation towards ectoderm
- (F) Bar graph of percentage of cells expressing *Pax6*, *Gata6* and *Nanog* transcripts after 12, 24 and 48 hours of towards ectoderm
- (G) Box and Whisker plots to show distribution of transcripts expression levels for *Pax6*, *Gata6* and *Nanog* after 12h, 24h and 48h differentiation towards ectoderm

Figure 3 – Lineage regulators are required for the downregulation of pluripotency regulators

- (A) Representative images of *Sox2*, *Oct4*, *Nanog* and *Esrrb* expression after 96 hours of differentiation with RA for *Pax6*(-/-), *Gata6*(-/-) and WT cells.
- (B) Box and whisker plots for quantification of *Sox2*, *Oct4*, *Nanog* and *Esrrb* expression of *Pax6*(-/-) and *Gata6*(-/-) cells after 96 hours of differentiation compared to WT cells.
- (C) Bar plots showing percentage of cells for WT, *Pax6*(-/-) and *Gata6*(-/-) expression more than 10 transcripts of *Oct4* and *Sox2* as well more than 5 transcripts of *Nanog* and *Esrrb* after 96 hours of differentiation.
- (D) Representative images of *Sox2*, *Oct4* and *Nanog* expression in WT, *Pax6*(-/-) and *Gata6*(-/-) cell lines after 96 hours of differentiation with RA and replating in Serum and LIF conditions for an additional 5 days. After 5 days in Serum and LIF the cells were harvested, fixed and imaged with fish.
- (E) Box and whisker plots for quantification of *Sox2*, *Oct4*, *Nanog* and *Esrrb* expression

of cells replated in Serum and LIF after 96 hours of differentiation.

Figure S4 – Related to Figure 3

- (A) Sequencing results for Pax6(-/-). The locus was PCR amplified and TOPO cloned. 15 Topo clones were sequenced and the two alleles (identical 8bp deletion), resulting in a homozygous frameshift are shown.
- (B) Sequencing results for Gata6(-/-). The locus was PCR amplified and TOPO cloned. Topo clones were sequenced and the two alleles, 1 allele showed a 25bp deletion while the other showed a 20bp deletion, resulting in a homozygous frameshift
- (C) Representative merged images for Sox2, Oct4 in 2i for WT, Pax6(-/-) and Gata6(-/-) cell line and Box and whisker plots for quantification of Sox2, Oct4 in WT and KO cell lines
- (D) Representative Images for Pax6 and Gata6 expression in WT, Pax6(-/-) and Gata6(-/-) cells after 96 hours RA differentiation

Fig. 4 – Lineage regulators are required for the exit of pluripotency in multiple differentiation systems

- (A) Representative images of Sox2, Oct4, Nanog and Esrrb expression after 96 hours of differentiation for Pax6(-/-) and WT cells differentiated towards ectoderm.
- (B) Box and whisker plots for quantification of Sox2, Oct4, Nanog and Esrrb expression in WT and Pax6(-/-) cells after 96 hours of differentiation towards ectoderm.
- (C) Bar plots showing percentage of cells for WT cell and Pax6(-/-) expression more than 10 transcripts of Oct4 and Sox2 as well more than 5 transcripts of Nanog and Esrrb after 96 hours of differentiation towards ectoderm.
- (D) Colony formation assay to determine pluripotency status of cells. Cells were exposed to various durations of RA and then replated in 2i. After 48 hours the surviving colonies were quantified.
- (E) Box and whisker plots comparing levels of Pluripotency factors (Oct4, Sox2, Nanog and Esrrb) in Pax6(-/-) and Gata6 (-/-) differentiation with towards either only ectoderm or endoderm and ectoderm after 96 hours (see Supplementary Figure 2A,B and Supplementary Figure 3 A-D)

Fig. 5 – The mode of exiting pluripotency is lineage specific

- (A) Representative sm-Fish images For Sox2 and Oct4 in WT cells under 2i conditions.
- (B) Scatter plots for quantification of Oct4 and Sox2 expression. Subpopulations were calculated through a ratio analysis of Sox2 and Oct4 within a single cell (Supplementary Methods). Cells expressing high level oct4 and low sox2 are shown in blue. Cells with a stable ratio of Oct4 and Sox2 are shown in grey. Cells expressing low oct4 but high sox2 are shown in pink
- (C) Representative smFISH images at 24h for Sox2 and Oct4
- (D) Scatter plots for quantification of Oct4 and Sox2 expression at 24h.
- (E) Representative smFISH images at 72h for Sox2 and Oct4
- (F) Scatter plots for quantification of Oct4 and Sox2 expression at 72h.
- (G) Representative images at 24h and 72h showing expression of Oct4, Sox2 and Gata6.
- (H) Box and Whisker plots showing Gata6 high and Gata6 low populations (see

Extendend Experimental Methods) and quantification of Sox2 and Oct4 in the two populations.

- (i) Box and Whisker plots for the quantification of Gata6 transcript levels in the Oct4high/Sox2low (Blue), Stable ratio (grey) and Oct4low/Sox2 high (pink) at 24h and 72h.

Fig. S5 – Related to Figure 5 (The mode of exiting pluripotency is lineage specific)

- (A) Representative images at 24h and 72h showing expression of Oct4, Sox2 and Pax6
- (B) Box and Whisker plot for the quantification of Pax6 transcript levels in the different subpopulations at 24h and 72h. (see Fig. 5I)
- (C) Quantification of Sox2 and Oct4 levels at 72 hours in Pax6high and Pax6low population (see Fig. 5H)

Fig. 6 – The lineage specific mode of exiting out of pluripotency is dependent on the presence of lineage regulators

- (A) Representative sm-Fish images For Sox2 and Oct4 in Gata6(-/-) cells after 96 hours of differentiation
- (B) Scatter plots for Gata6(-/-) Sox2 and Oct4 for subpopulation analysis (see Fig.5B)
- (C) Representative sm-Fish images For Sox2 and Oct4 in Pax6(-/-) cells after 96 hours of differentiation
- (D) Scatter plots for Pax6(-/-) Sox2 and Oct4 for subpopulation analysis (see Fig.5B)
- (E) Bar plot for the quantification of subpopulations at 24h and 72h of differentiation for WT, Pax6(-/-) and Gata6(-/-) cells

References

- Aksoy, I., Jauch, R., Chen, J., Dyla, M., Divakar, U., Bogu, G.K., Teo, R., Leng Ng, C.K., Herath, W., Lili, S., et al. (2013). Oct4 switches partnering from Sox2 to Sox17 to reinterpret the enhancer code and specify endoderm. *EMBO J* 32, 938-953.
- Ben-David, U., and Benvenisty, N. (2011). The tumorigenicity of human embryonic and induced pluripotent stem cells. *Nat Rev Cancer* 11, 268-277.
- Betschinger, J., Nichols, J., Dietmann, S., Corrin, P.D., Paddison, P.J., and Smith, A. (2013). Exit from pluripotency is gated by intracellular redistribution of the bHLH transcription factor Tfe3. *Cell* 153, 335-347.
- Boroviak, T., and Nichols, J. (2014). The birth of embryonic pluripotency. *Philos Trans R Soc Lond B Biol Sci* 369.
- Boyer, L.A., Lee, T.I., Cole, M.F., Johnstone, S.E., Levine, S.S., Zucker, J.P., Guenther, M.G., Kumar, R.M., Murray, H.L., Jenner, R.G., et al. (2005). Core transcriptional regulatory circuitry in human embryonic stem cells. *Cell* 122, 947-956.

Buganim, Y., Faddah, D.A., and Jaenisch, R. (2013). Mechanisms and models of somatic cell reprogramming. *Nat Rev Genet* 14, 427-439.

Capo-Chichi, C.D., Rula, M.E., Smedberg, J.L., Vanderveer, L., Parmacek, M.S., Morrissey, E.E., Godwin, A.K., and Xu, X.X. (2005). Perception of differentiation cues by GATA factors in primitive endoderm lineage determination of mouse embryonic stem cells. *Dev Biol* 286, 574-586.

Chambers, I., and Tomlinson, S.R. (2009). The transcriptional foundation of pluripotency. *Development* 136, 2311-2322.

Chen, X., Xu, H., Yuan, P., Fang, F., Huss, M., Vega, V.B., Wong, E., Orlov, Y.L., Zhang, W., Jiang, J., et al. (2008). Integration of external signaling pathways with the core transcriptional network in embryonic stem cells. *Cell* 133, 1106-1117.

Cho, K.W., Kwon, H.J., Shin, J.O., Lee, J.M., Cho, S.W., Tickle, C., and Jung, H.S. (2012). Retinoic acid signaling and the initiation of mammary gland development. *Dev Biol* 365, 259-266. D'Aniello, E., and Waxman, J.S. (2015). Input overload: Contributions of retinoic acid signaling feedback mechanisms to heart development and teratogenesis. *Dev Dyn* 244, 513-523.

De Los Angeles, A., Ferrari, F., Xi, R., Fujiwara, Y., Benvenisty, N., Deng, H., Hochedlinger, K., Jaenisch, R., Lee, S., Leitch, H.G., et al. (2015). Hallmarks of pluripotency. *Nature* 525, 469-478. Dunn, S.J., Martello, G., Yordanov, B., Emmott, S., and Smith, A.G. (2014). Defining an essential transcription factor program for naive pluripotency. *Science* 344, 1156-1160.

Ebert, A.D., Liang, P., and Wu, J.C. (2012). Induced pluripotent stem cells as a disease modeling and drug screening platform. *J Cardiovasc Pharmacol* 60, 408-416.

Faddah, D.A., Wang, H., Cheng, A.W., Katz, Y., Buganim, Y., and Jaenisch, R. (2013). Single-cell analysis reveals that expression of nanog is biallelic and equally variable as that of other pluripotency factors in mouse ESCs. *Cell Stem Cell* 13, 23-29.

Fagnocchi, L., Cherubini, A., Hatsuda, H., Fasciani, A., Mazzoleni, S., Poli, V., Berno, V., Rossi, R.L., Reinbold, R., Ende, M., et al. (2016). A Myc-driven self-reinforcing regulatory network maintains mouse embryonic stem cell identity. *Nat Commun* 7, 11903.

Festuccia, N., Osorno, R., Wilson, V., and Chambers, I. (2013). The role of pluripotency gene regulatory network components in mediating transitions between pluripotent cell states. *Curr Opin Genet Dev* 23, 504-511.

Filipczyk, A., Marr, C., Hastreiter, S., Feigelman, J., Schwarzfischer, M., Hoppe, P.S., Loeffler, D., Kokkaliaris, K.D., Ende, M., Schauburger, B., et al. (2015). Network plasticity of pluripotency transcription factors in embryonic stem cells. *Nat Cell Biol* 17, 1235-1246.

- Goolam, M., Scialdone, A., Graham, S.J., Macaulay, I.C., Jedrusik, A., Hupalowska, A., Voet, T., Marioni, J.C., and Zernicka-Goetz, M. (2016). Heterogeneity in Oct4 and Sox2 Targets Biases Cell Fate in 4-Cell Mouse Embryos. *Cell* 165, 61-74.
- Hansen, C.H., and van Oudenaarden, A. (2013). Allele-specific detection of single mRNA molecules in situ. *Nat Methods* 10, 869-871.
- Hayashi, K., Ohta, H., Kurimoto, K., Aramaki, S., and Saitou, M. (2011). Reconstitution of the mouse germ cell specification pathway in culture by pluripotent stem cells. *Cell* 146, 519-532.
- Huang, G., Ye, S., Zhou, X., Liu, D., and Ying, Q.L. (2015). Molecular basis of embryonic stem cell self-renewal: from signaling pathways to pluripotency network. *Cell Mol Life Sci* 72, 1741- 1757.
- Jaenisch, R., and Young, R. (2008). Stem cells, the molecular circuitry of pluripotency and nuclear reprogramming. *Cell* 132, 567-582.
- Kalkan, T., and Smith, A. (2014). Mapping the route from naive pluripotency to lineage specification. *Philos Trans R Soc Lond B Biol Sci* 369.
- Kawamata, S., Kanemura, H., Sakai, N., Takahashi, M., and Go, M.J. (2015). Design of a Tumorigenicity Test for Induced Pluripotent Stem Cell (iPSC)-Derived Cell Products. *J Clin Med* 4, 159-171.
- Kopp, J.L., Ormsbee, B.D., Desler, M., and Rizzino, A. (2008). Small increases in the level of Sox2 trigger the differentiation of mouse embryonic stem cells. *Stem Cells* 26, 903-911.
- Kumar, R.M., Cahan, P., Shalek, A.K., Satija, R., DaleyKeyser, A.J., Li, H., Zhang, J., Pardee, K., Gennert, D., Trombetta, J.J., et al. (2014). Deconstructing transcriptional heterogeneity in pluripotent stem cells. *Nature* 516, 56-61.
- Lawson, K.A., Meneses, J.J., and Pedersen, R.A. (1991). Clonal analysis of epiblast fate during germ layer formation in the mouse embryo. *Development* 113, 891-911.
- Leeb, M., Dietmann, S., Paramor, M., Niwa, H., and Smith, A. (2014). Genetic exploration of the exit from self-renewal using haploid embryonic stem cells. *Cell Stem Cell* 14, 385-393.
- Lodato, M.A., Ng, C.W., Wamstad, J.A., Cheng, A.W., Thai, K.K., Fraenkel, E., Jaenisch, R., and Boyer, L.A. (2013). SOX2 co-occupies distal enhancer elements with distinct POU factors in ESCs and NPCs to specify cell state. *PLoS Genet* 9, e1003288.
- Loh, K.M., Lim, B., and Ang, L.T. (2015). Ex uno plures: molecular designs for embryonic pluripotency. *Physiol Rev* 95, 245-295.
- Mammadova, A., Zhou, H., Carels, C.E., and Von den Hoff, J.W. (2016). Retinoic acid

signalling in the development of the epidermis, the limbs and the secondary palate. *Differentiation*.

Montserrat, N., Nivet, E., Sancho-Martinez, I., Hishida, T., Kumar, S., Miquel, L., Cortina, C., Hishida, Y., Xia, Y., Esteban, C.R., et al. (2013). Reprogramming of human fibroblasts to pluripotency with lineage specifiers. *Cell Stem Cell* 13, 341-350.

Niwa, H., Miyazaki, J., and Smith, A.G. (2000). Quantitative expression of Oct-3/4 defines differentiation, dedifferentiation or self-renewal of ES cells. *Nat Genet* 24, 372-376

Olsson, A., Venkatasubramanian, M., Chaudhri, V.K., Aronow, B.J., Salomonis, N., Singh, H., and Grimes, H.L. (2016). Single-cell analysis of mixed-lineage states leading to a binary cell fate choice. *Nature* 537, 698-702.

Orkin, S.H., Wang, J., Kim, J., Chu, J., Rao, S., Theunissen, T.W., Shen, X., and Levasseur, D.N. (2008). The transcriptional network controlling pluripotency in ES cells. *Cold Spring Harb Symp Quant Biol* 73, 195-202.

Peterson, S.E., Garitaonandia, I., and Loring, J.F. (2016). The tumorigenic potential of pluripotent stem cells: What can we do to minimize it? *Bioessays* 38 Suppl 1, S86-95.

Raj, A., van den Bogaard, P., Rifkin, S.A., van Oudenaarden, A., and Tyagi, S. (2008). Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods* 5, 877-879.

Rhinn, M., and Dolle, P. (2012). Retinoic acid signalling during development. *Development* 139, 843-858.

Saha, K., and Jaenisch, R. (2009). Technical challenges in using human induced pluripotent stem cells to model disease. *Cell Stem Cell* 5, 584-595.

Schlupf, J., and Steinbeisser, H. (2014). IGF antagonizes the Wnt/beta-Catenin pathway and promotes differentiation of extra-embryonic endoderm. *Differentiation* 87, 209-219.

Semrau, S., Crosetto, N., Bienko, M., Boni, M., Bernasconi, P., Chiarle, R., and van Oudenaarden, A. (2014). FuseFISH: robust detection of transcribed gene fusions in single cells. *Cell Rep* 6, 18-23.

Shimozaki, K., Nakashima, K., Niwa, H., and Taga, T. (2003). Involvement of Oct3/4 in the enhancement of neuronal differentiation of ES cells in neurogenesis-inducing cultures. *Development* 130, 2505-2512.

Shu, J., and Deng, H. (2013). Lineage specifiers: new players in the induction of pluripotency. *Genomics Proteomics Bioinformatics* 11, 259-263.

Shu, J., Wu, C., Wu, Y., Li, Z., Shao, S., Zhao, W., Tang, X., Yang, H., Shen, L., Zuo, X., et al. (2013). Induction of pluripotency in mouse somatic cells with lineage specifiers.

Cell 153, 963- 975.

Silva, J., Barrandon, O., Nichols, J., Kawaguchi, J., Theunissen, T.W., and Smith, A. (2008). Promotion of reprogramming to ground state pluripotency by signal inhibition. *PLoS Biol* 6, e253. Svendsen, C.N. (2013). Back to the future: how human induced pluripotent stem cells will transform regenerative medicine. *Hum Mol Genet* 22, R32-38.

Tabar, V., and Studer, L. (2014). Pluripotent stem cells in regenerative medicine: challenges and recent progress. *Nat Rev Genet* 15, 82-92.

Takahashi, K., and Yamanaka, S. (2006). Induction of pluripotent stem cells from mouse embryonic and adult fibroblast cultures by defined factors. *Cell* 126, 663-676.

Thomson, M., Liu, S.J., Zou, L.N., Smith, Z., Meissner, A., and Ramanathan, S. (2011). Pluripotency factors in embryonic stem cells regulate differentiation into germ layers. *Cell* 145, 875-889.

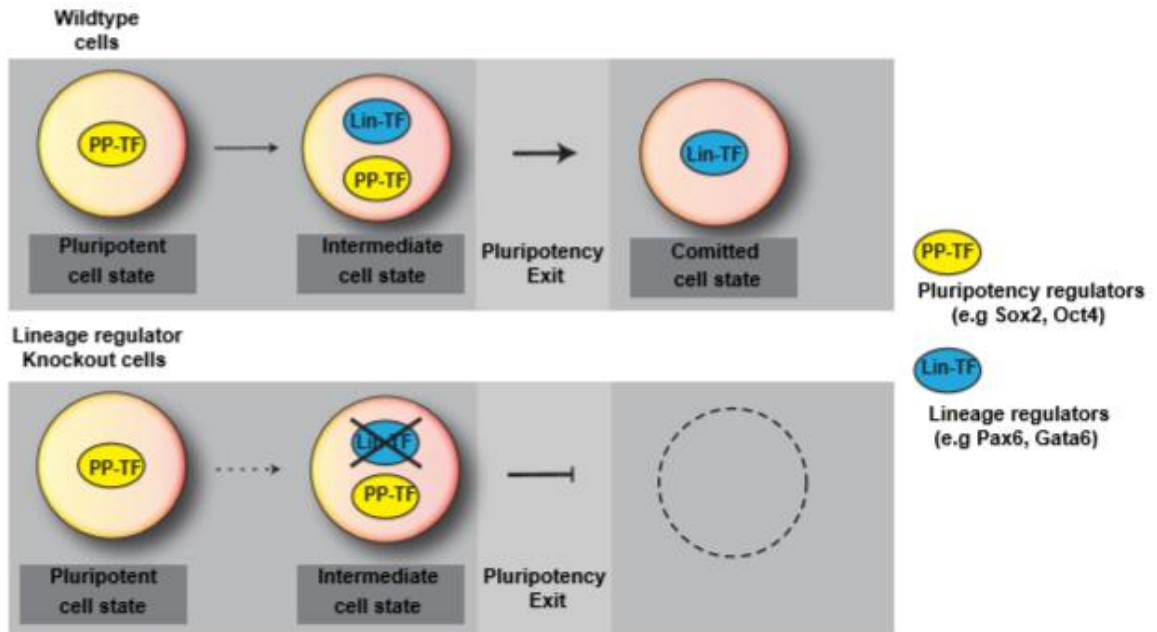
Wang, J., Rao, S., Chu, J., Shen, X., Levasseur, D.N., Theunissen, T.W., and Orkin, S.H. (2006). A protein interaction network for pluripotency of embryonic stem cells. *Nature* 444, 364-368.

Yang, H., Wang, H., Shivalila, C.S., Cheng, A.W., Shi, L., and Jaenisch, R. (2013). One-step generation of mice carrying reporter and conditional alleles by CRISPR/Cas-mediated genome engineering. *Cell* 154, 1370-1379.

Yang, S.H., Kalkan, T., Morrisroe, C., Smith, A., and Sharrocks, A.D. (2012). A genome-wide RNAi screen reveals MAP kinase phosphatases as key ERK pathway regulators during embryonic stem cell differentiation. *PLoS Genet* 8, e1003112.

Ying, Q.L., Wray, J., Nichols, J., Battle-Morera, L., Doble, B., Woodgett, J., Cohen, P., and Smith, A. (2008). The ground state of embryonic stem cell self-renewal. *Nature* 453, 519-523.

Zhang, X., Huang, C.T., Chen, J., Pankratz, M.T., Xi, J., Li, J., Yang, Y., Lavaute, T.M., Li, X.J., Ayala, M., et al. (2010). Pax6 is a human neuroectoderm cell fate determinant. *Cell Stem Cell* 7, 90-100.



Graphical Abstract

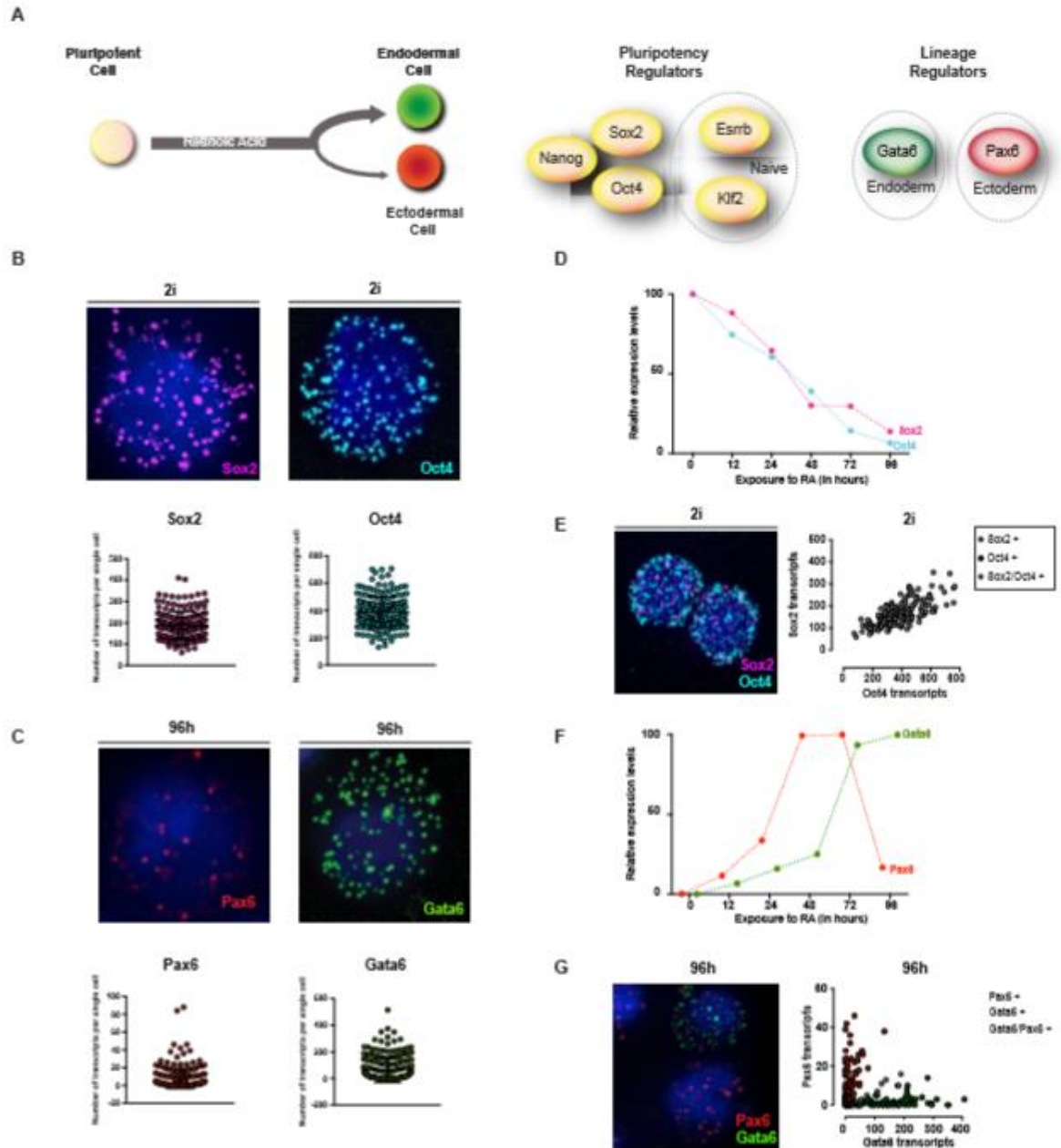


Figure 1. Single molecule RNA Fish based detection system to study the exit from pluripotency at a single cell resolution

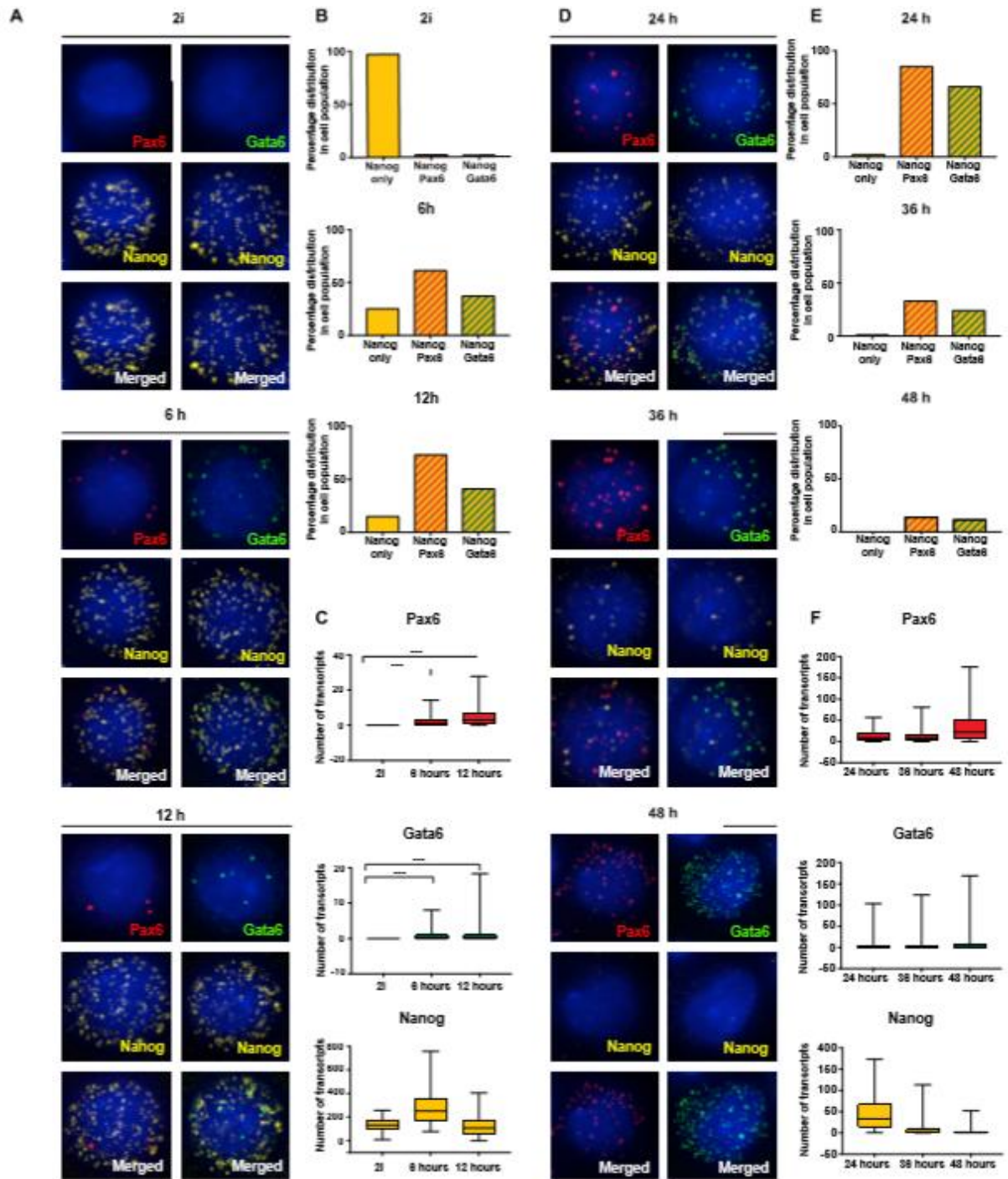


Figure 2 - Lineage regulators and pluripotency regulators are co-expressed prior and during the exit from pluripotency

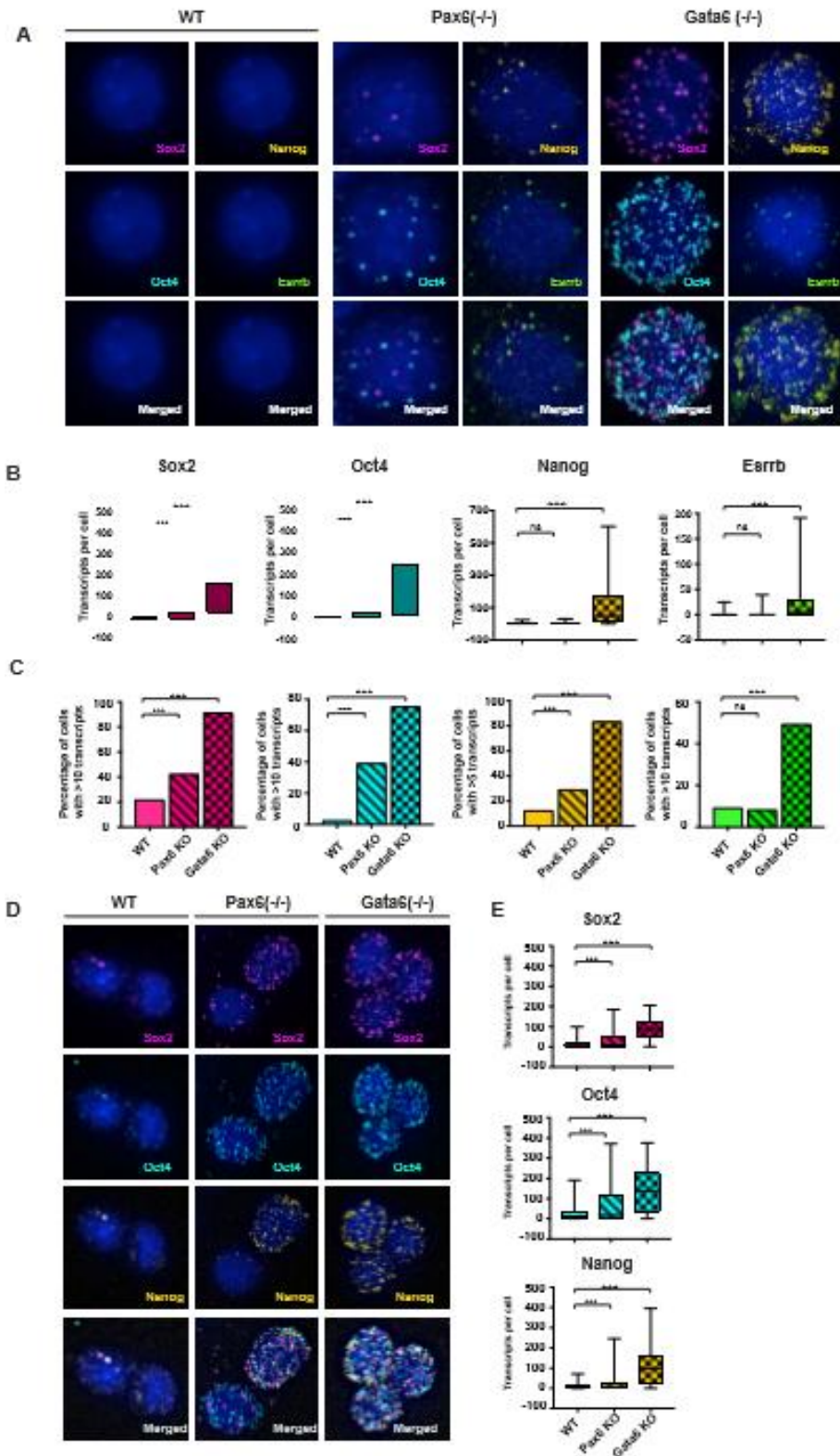


Figure 3. Lineage regulators are required for the downregulation of pluripotency regulators

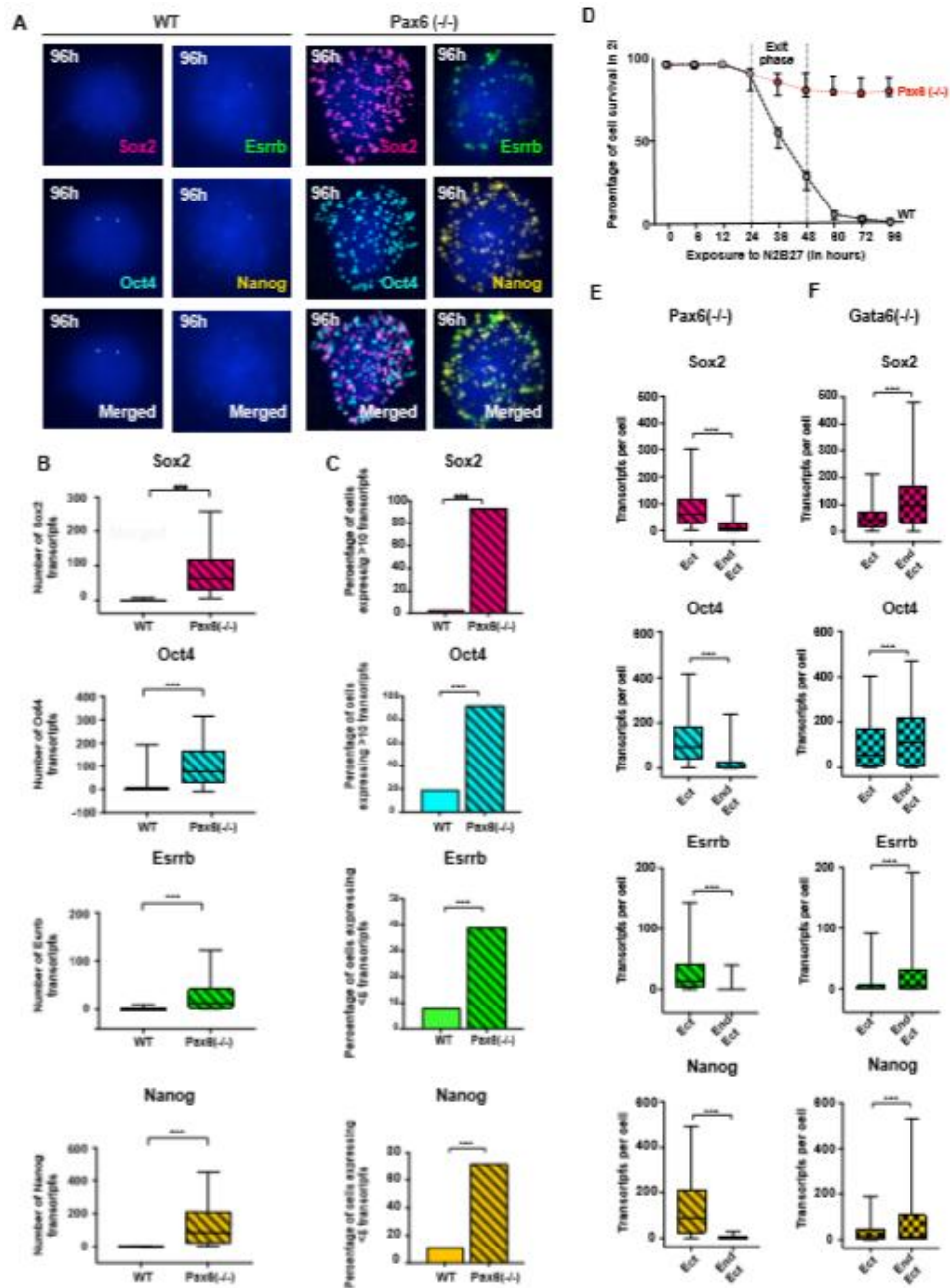


Figure 4. Lineage regulators are required for exit of pluripotency in multiple differentiation systems

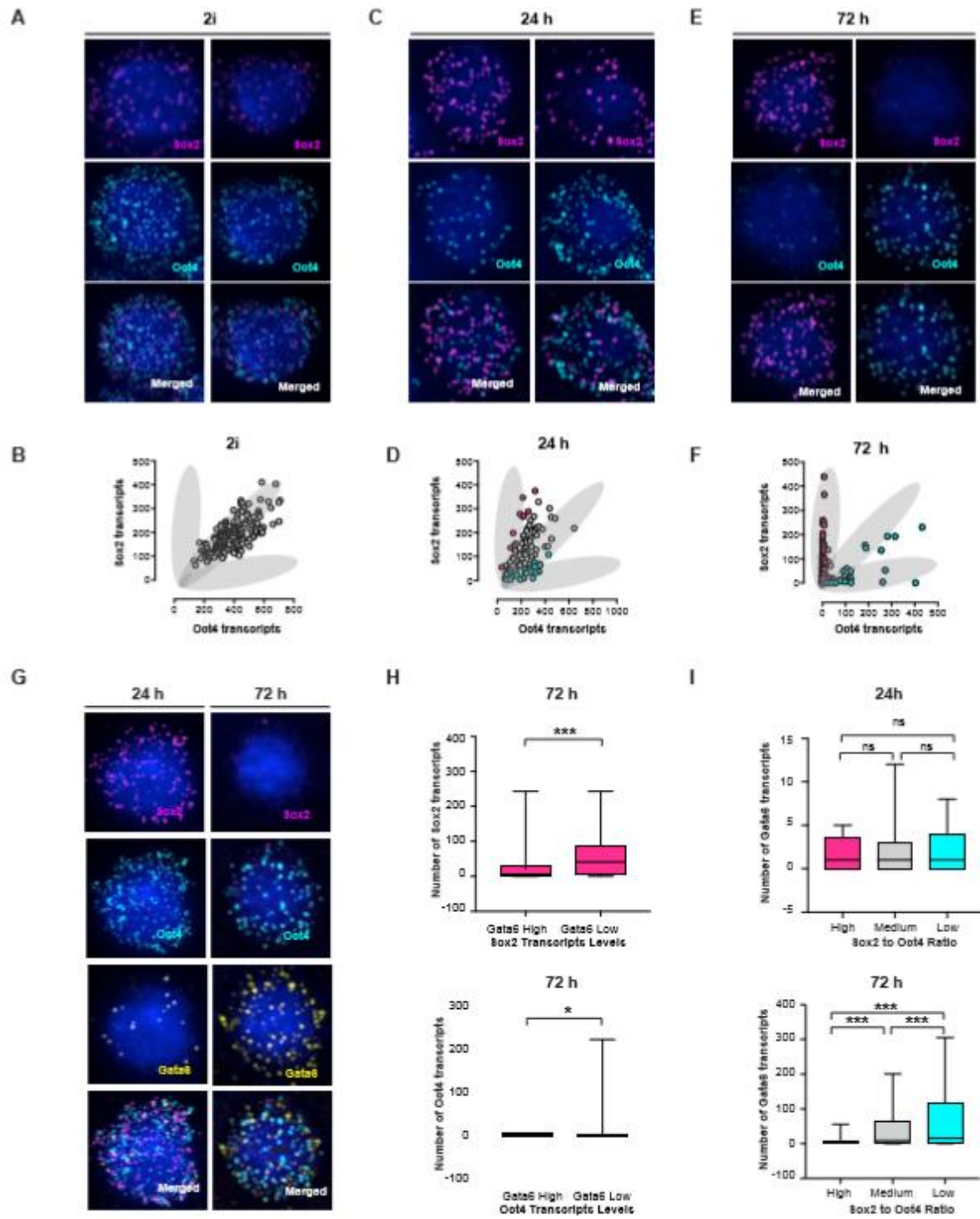


Figure 5. The mode of exiting pluripotency is lineage specific

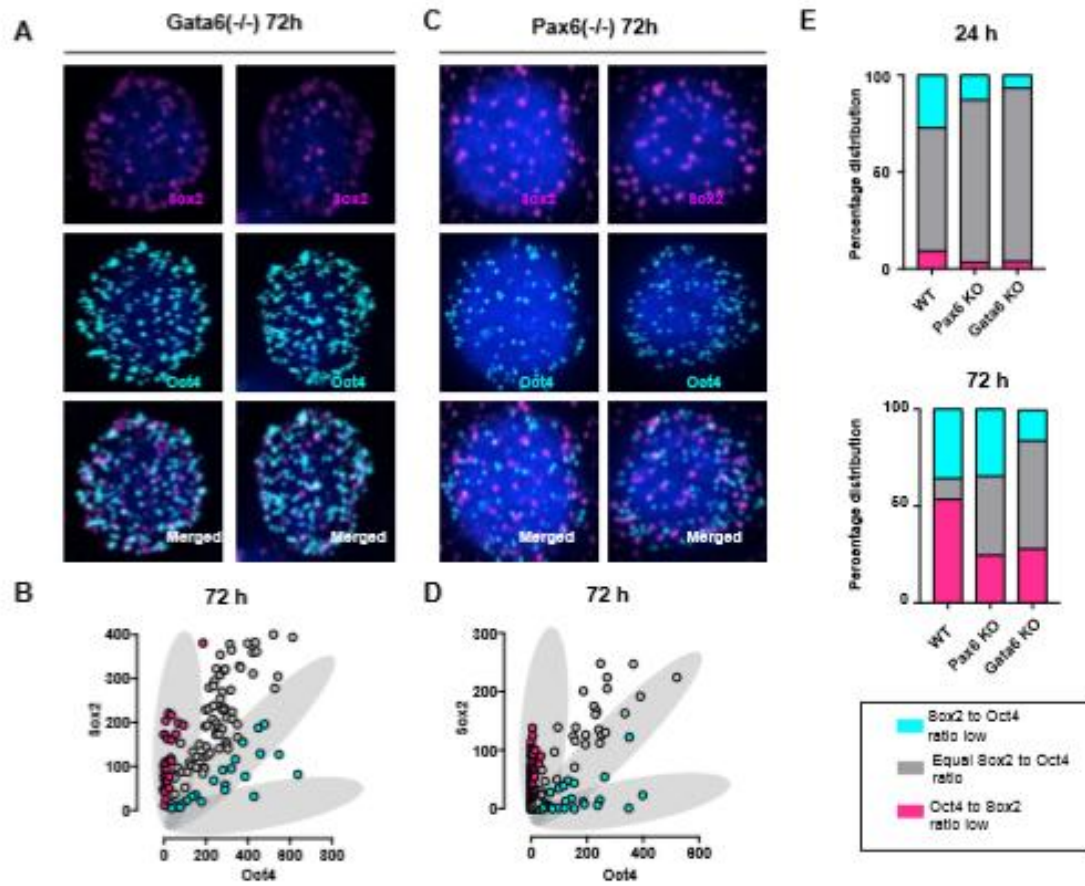


Figure 6. The lineage specific mode of exiting out pluripotency is dependent on the expression of lineage regulators

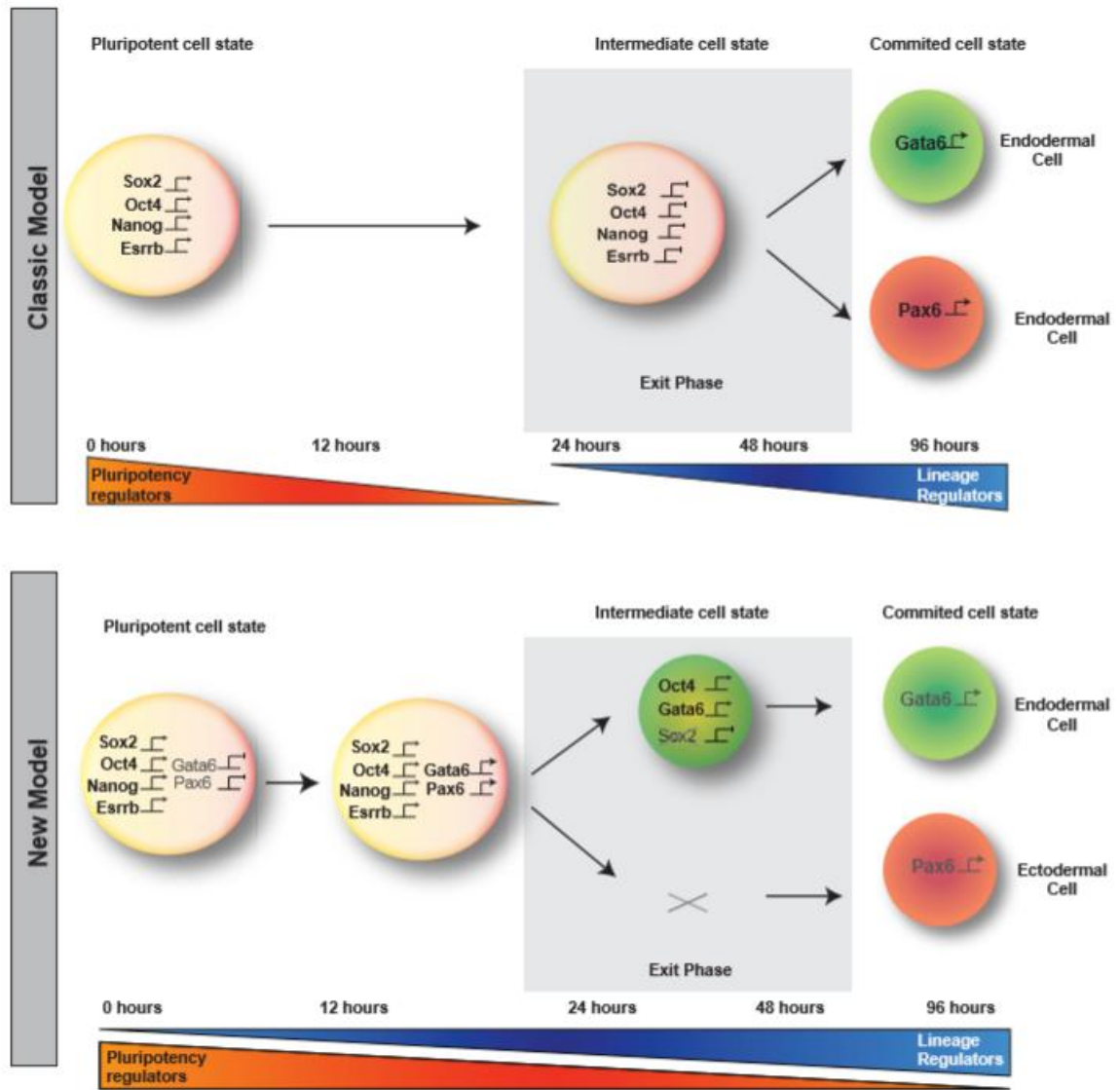


Figure 7. Pluripotency and lineage regulators form an interconnected network to control cell state transitions out of pluripotency

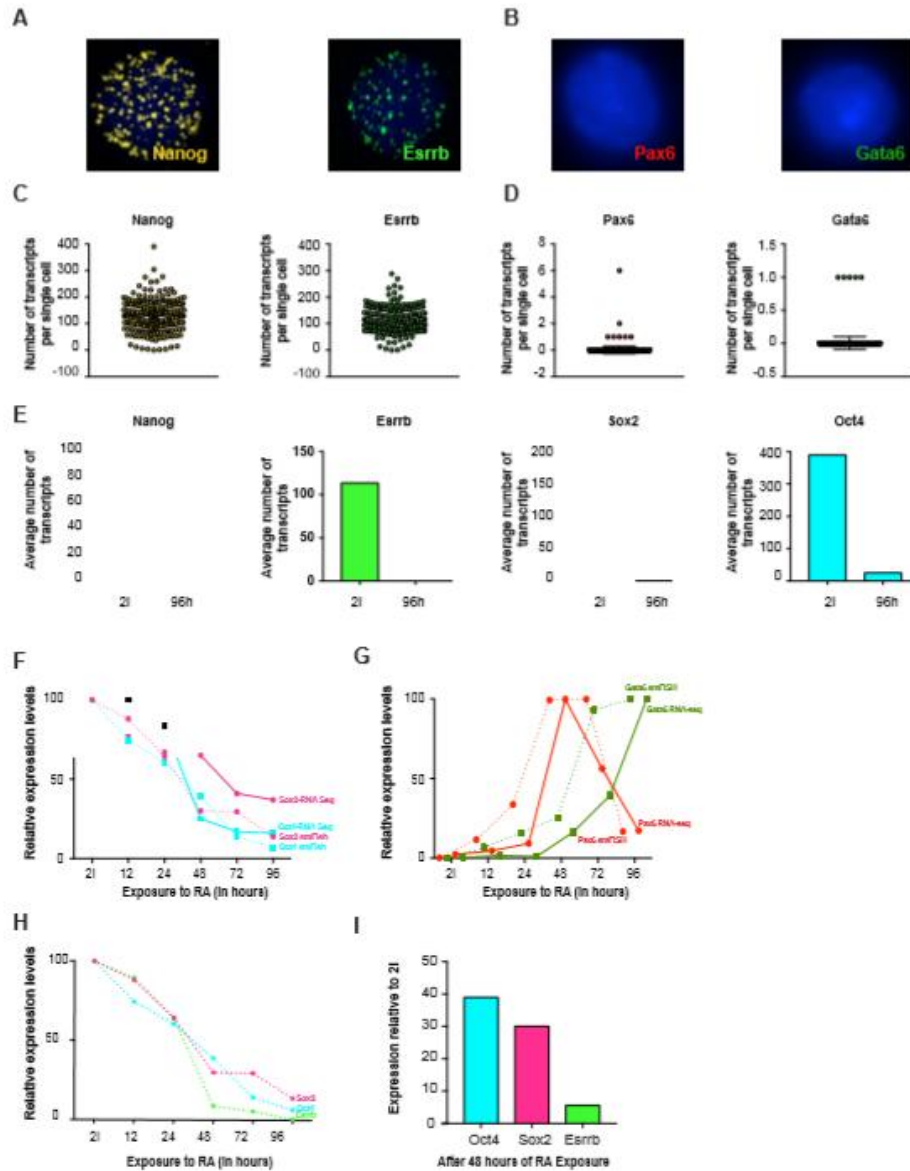


Figure S1. Related to Figure 1 (Single molecule RNA Fish based detection system to study the Exit from pluripotency at a single cell resolution)

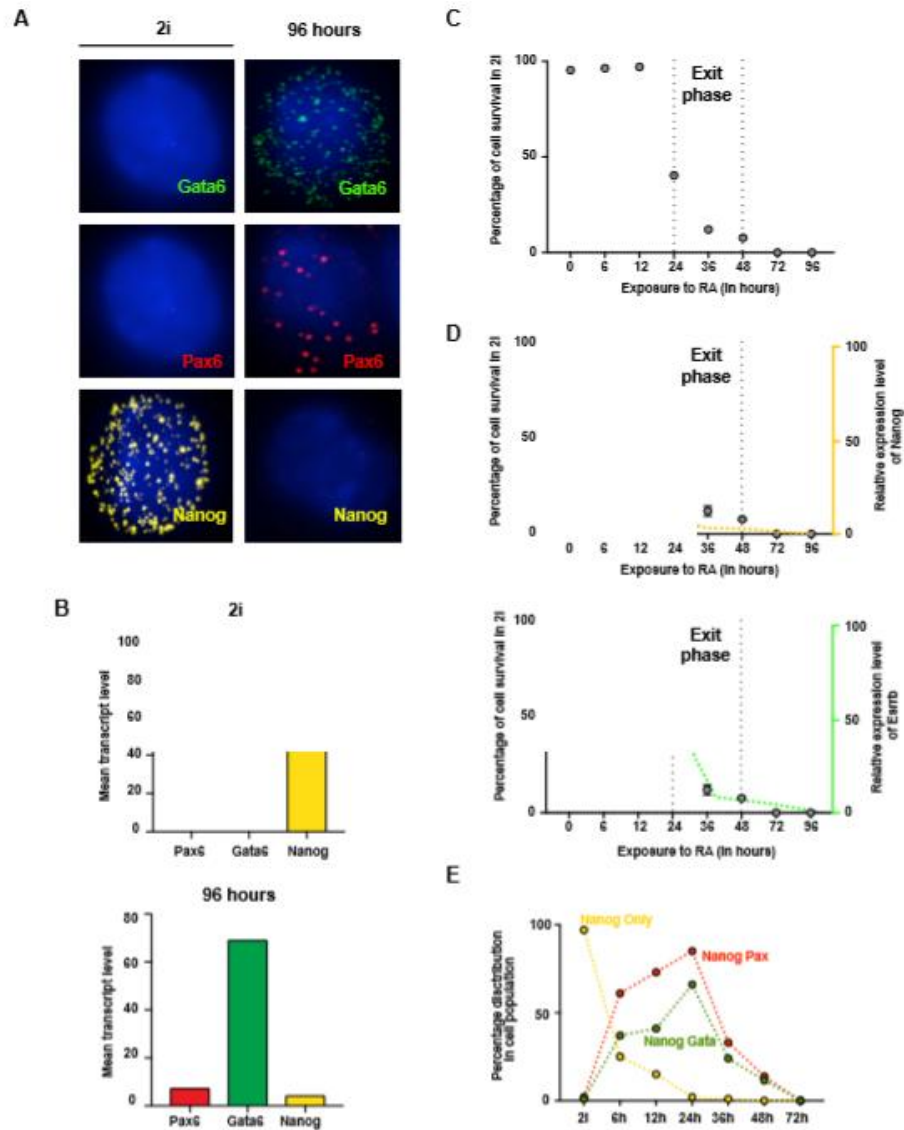
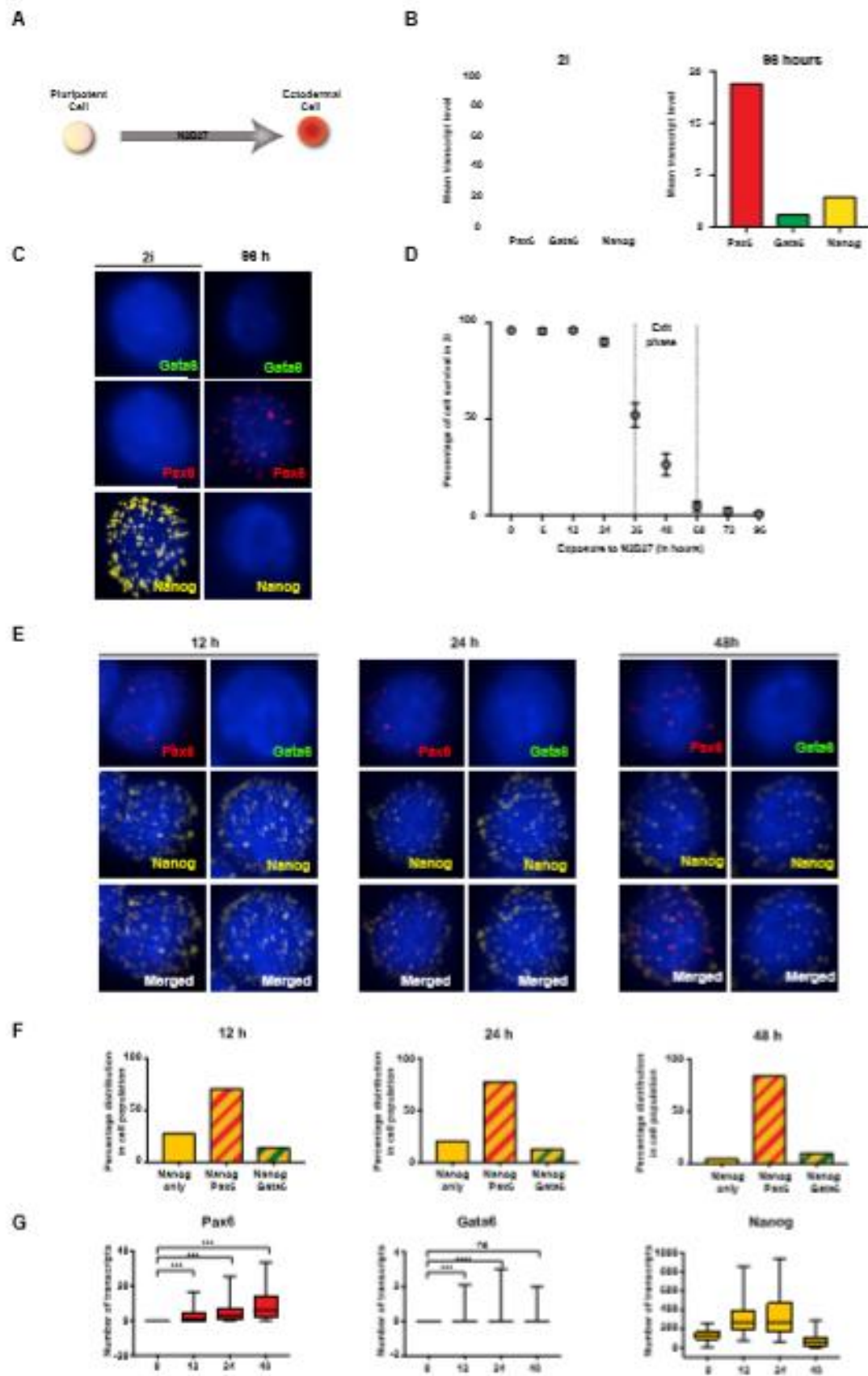


Figure S2. Related to Figure 2 (Lineage and pluripotency regulators are co-expressed prior and during the exit from pluripotency)



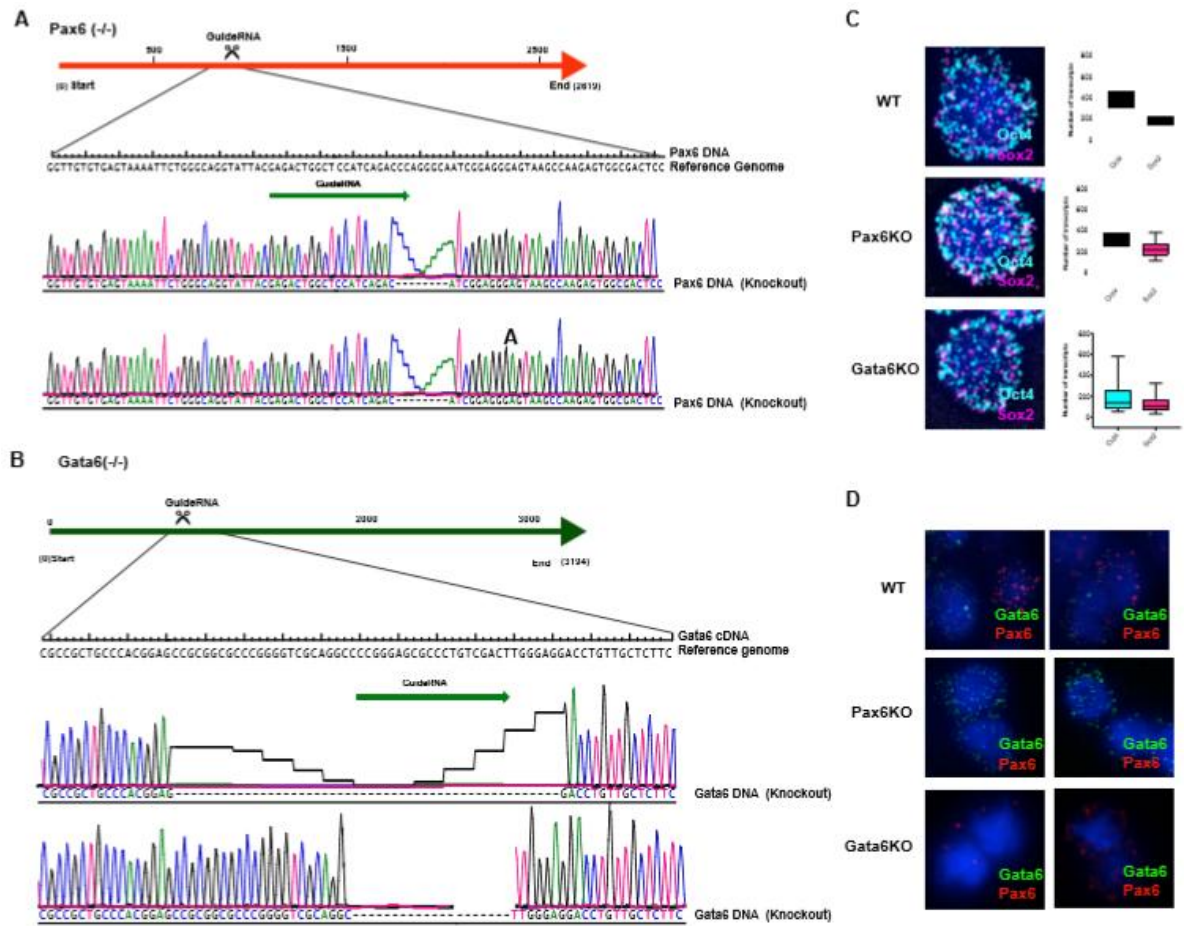


Figure S4 - Related to Figure 3 (Lineage regulators are required for the downregulation of pluripotency regulators)

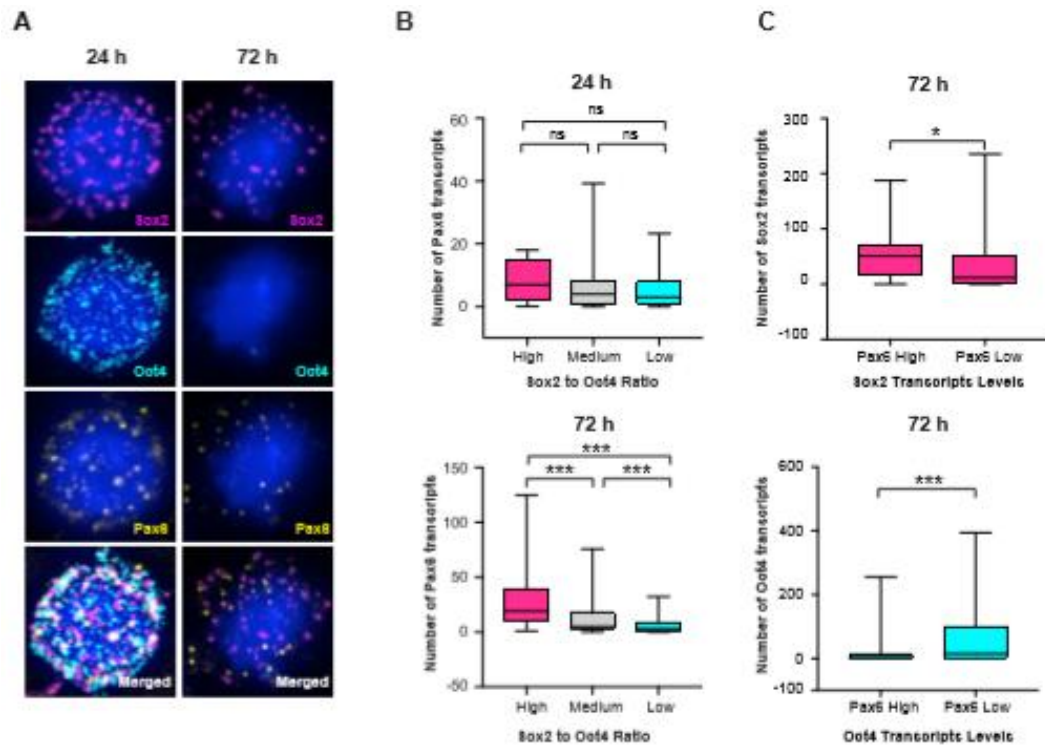

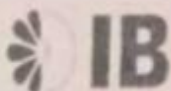


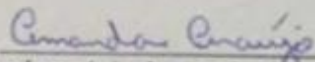
Fig S5. Related to Figure 5 (The mode of exiting pluripotency is lineage specific)

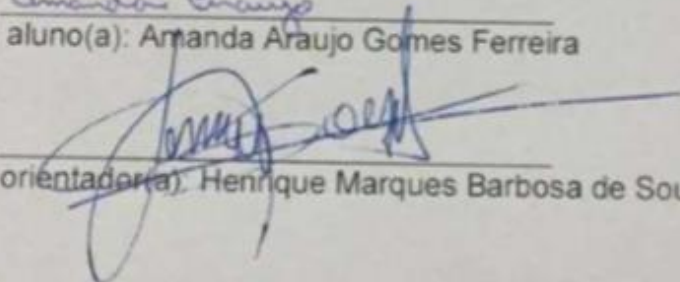
Anexo 2

 UNICAMP	<small>COORDENADORIA DE PÓS-GRADUAÇÃO INSTITUTO DE BIOLOGIA Universidade Estadual de Campinas Caixa Postal 6109, 13083-970, Campinas, SP, Brasil Fone (19) 3521-6378. email: cpqib@unicamp.br</small>	 IB
--	---	---

DECLARAÇÃO

Em observância ao **§5º do Artigo 1º da Informação CCPG-UNICAMP/001/15**, referente a Bioética e Biossegurança, declaro que o conteúdo de minha Dissertação de Mestrado, intitulada ***"O papel de Coup-TFII na aquisição de destinos celulares durante a diferenciação de Células-tronco embrionárias"***, desenvolvida no Programa de Pós-Graduação em Genética e Biologia Molecular do Instituto de Biologia da Unicamp, não versa sobre pesquisa envolvendo seres humanos, animais ou temas afetos a Biossegurança.

Assinatura: 
Nome do(a) aluno(a): **Amanda Araujo Gomes Ferreira**

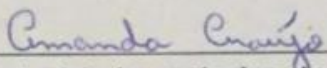
Assinatura: 
Nome do(a) orientador(a): **Henrique Marques Barbosa de Souza**

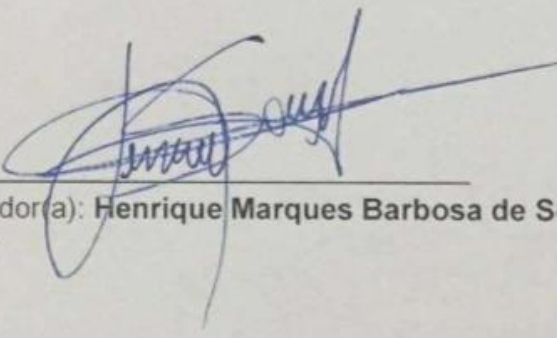
Data: 05/05/2017

Anexo 3**Declaração**

As cópias de artigos de minha autoria ou de minha co-autoria, já publicados ou submetidos para publicação em revistas científicas ou anais de congressos sujeitos a arbitragem, que constam da minha Dissertação/Tese de Mestrado/Doutorado, intitulada **O papel de Coup-TFII na aquisição de destinos celulares durante a diferenciação de Células-tronco embrionárias**, não infringem os dispositivos da Lei n.º 9.610/98, nem o direito autoral de qualquer editora.

Campinas, 05 de Maio de 2017

Assinatura : 
Nome do(a) autor(a): **Amanda Araújo Gomes Ferreira**
RG n.º 49561912-7

Assinatura : 
Nome do(a) orientador(a): **Henrique Marques Barbosa de Souza**
RG n.º 2.153.826-4